Social Science
Library
JUL 8 - 1

Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
AIR FORCE PERSONNEL AND TRAINING RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Cletus J. Burke, *Indiana University*
Robert M. Gagné, *AF Personnel and Training Research Center*
W. R. Garner, *The Johns Hopkins University*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *AF Personnel and Training Research Center*
Arthur L. Irion, *Tulane University*
Donald B. Lindsley, *University of California, Los Angeles*
Neal E. Miller, *Yale University*
Edwin B. Newman, *Harvard University*
Kenneth W. Spence, *State University of Iowa*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Lorraine Bouthilet, Managing Editor

CONTENTS

The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement: P. M. FITTS 381
The Relation of Different Levels and Kinds of Motivation to Variability of Behavior: R. L. DE VALOIS 392
An Experimental Distinction between Perceptual Process and Verbal Response: U. NEISSER 399
Eye-Hand Span in Simple Serial Tasks: E. C. POULTON 403
Transfer of Training as a Function of Degree of Response Overlearning: G. MANDLER ... 411
Resistance to Extinction as a Function of Variations in Stimuli Associated with Shock: H. MOLTZ 418
Temporal Numerosity: III. *Auditory Perception of Number:* P. G. CHEATHAM AND C. T. WHITE 425
Extinction of Instrumental Behavior as a Function of Frustation at Various Distances from the Goal: H. M. ADELMAN AND G. ROSENBAUM 429
Interpolation Accuracy as a Function of Visual Angle between Scale Marks: C. A. BAKER 433
Conditioned Stimulus Intensity and Response Speed: R. M. BRAGIEL AND C. C. PERKINS, JR. 437
Distributed Practice and Rote Learning in Concept Formation: J. RICHARDSON AND B. O. BERGUM 442
Visual-Motor Performance as a Function of Short-Duration Ambient Temperature: W. H. TEICHNER AND R. F. WEHRKAMP 447
Comparison of the Masked Thresholds of a Simulated Moving and Stationary Auditory Signal: R. C. WILCOTT AND R. S. GALES 451
Errors in Transfer Following Learning with Understanding: Further Studies with Katona's Card-Trick Experiments: E. R. HILGARD, R. D. EDGREN, AND R. P. IRVINE 457
Learning a Rate of Movement: R. S. LINCOLN 465

American Psychological Association

Vol. 47 No. 6 June 1954

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc. The subscription rate per volume is $7.50, or $15.00 annually. Single copies are $1.50. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology. Normally, articles of a length exceeding 15 printed pages cannot be accepted. However, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1954 the publication cost is $20 per printed page. Authors of priority publications receive no free offprints.

Address all articles submitted for regular or priority publication to the editor, Arthur W. Melton, Air Force Personnel and Training Research Center, Lackland Air Force Base, San Antonio, Texas.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 (4, Suppl.), 389–449. In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Ss*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879.

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Copyright, 1954, by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 47, NO. 6 JUNE, 1954

THE INFORMATION CAPACITY OF THE HUMAN MOTOR SYSTEM IN CONTROLLING THE AMPLITUDE OF MOVEMENT[1]

PAUL M. FITTS[2]

The Ohio State University

Information theory has recently been employed to specify more precisely than has hitherto been possible man's capacity in certain sensory, perceptual, and perceptual-motor functions (**5, 10, 13, 15, 17, 18**). The experiments reported in the present paper extend the theory to the human motor system. The applicability of only the basic concepts, *amount of information*, *noise*, *channel capacity*, and *rate of information transmission*, will be examined at this time. General familiarity with these concepts as formulated by recent writers (**4, 11, 20, 22**) is assumed.

Strictly speaking, we cannot study man's motor system at the behavioral level in isolation from its associated sensory mechanisms. We can only analyze the behavior of the entire receptor-neural-effector system. However, by asking *S* to make rapid and uniform responses that have been highly overlearned, and by holding all relevant stimulus conditions constant with the exception of those resulting from *S*'s own movements, we can create an experimental situation in which it is reasonable to assume that performance is limited primarily by the capacity of the motor system. The motor system in the present case is defined as including the visual and proprioceptive feedback loops that permit *S* to monitor his own activity.

The information capacity of the motor system is specified by its ability to produce consistently one class of movement from among several alternative movement classes. The greater the number of alternative classes, the greater is the information capacity of a particular type of response. Since measurable aspects of motor responses, such as their force, direction, and amplitude, are continuous variables, their information capacity is limited only by the amount of statistical variability, or noise, that is characteristic of repeated efforts to produce the same response. The information capacity of the motor system, therefore, can be

[1] This research was supported in part by the United States Air Force under Contract No. AF 33(038)-10528 with the Ohio State University Research Foundation, monitored by the Air Force Personnel and Training Research Center. Permission is granted for reproduction, publication, use, and disposal in whole or in part by or for the United States Government.

[2] Charles Kelly, Robert Silverman, and Charlotte Christner assisted in collecting the data here reported.

inferred from measures of the variability of successive responses that S attempts to make uniform.

It is possible to determine experimentally the noise associated with each category of response amplitude and rate, and to infer the average information capacity per response and the maximum average rate of information transmission from the ratio of the magnitude of this noise to the magnitude of the possible range of responses. This line of reasoning agrees with Miller's (20) suggestion that the concept of information capacity can be interpreted as a sort of modern version of the traditional Weber-fraction and is consistent with Theorem 17 in Shannon (22).[3]

The present experiments are limited to motor tasks in which S is asked to make successive responses having a specified amplitude of movement. The information in which we are interested is generated in discrete increments, an increment being added by each successive "response." The relations to be studied are those holding among (*a*) the average amplitude, (*b*) the average duration, and (*c*) the distribution of successive amplitudes (variability) of a series of rapid movements that S attempts to produce with uniform amplitude. The thesis to be examined is that the channel capacity of the motor system, in a task involving a particular limb, a particular set of muscles, and a particular type of motor behavior (within limits), is independent of average amplitude and of specified permissible variability (movement tolerance).

The need for a unifying concept of motor capacity is indicated by the apparent difficulty of reconciling many of the facts reported in the literature on motor skill. These difficulties have in large measure stemmed from failure to recognize that the amplitude, the duration, and the variability of movements are interrelated.

Early investigators were interested chiefly in activities such as running, tapping with a telegraph key, writing longhand, and keeping time with a baton in which visual monitoring of successive movement is minimal and in which successive movements do not have to be terminated precisely. The duration of this type of movement is essentially independent of its extent. Bryan (2), for example, reported that tapping rate varied only from 4.6 to 6 cps over an amplitude range from 1 to 40 mm. Noting that his results confirmed earlier experiments by Von Kries, he concluded that when movements are made at maximum rate ". . . the rate varies slightly with different muscles; increases with practice; but is not affected by variations of the amplitude of motion within wide limits, a certain medium amplitude requiring less time than longer or shorter distances" (2, p. 138). Freeman (9) studied handwriting movement and found little variation in the total time required to write letters as their size was varied. Stetson and McDill state categorically that "since the duration of the rapid movement is fixed, there can be but two variables, extent and force" (23, p. 25), and Hartson (12) makes fixed duration the basic characteristic of the ballistic movement.

Recent investigators (1, 3, 14, 21), studying tasks that require S to exercise precise control of movement amplitude but rely on his subjective standard of acceptable movement precision, have reported results that are contrary to those cited in the preceding paragraph. An example of such a task is moving a control quickly so as to realign a spot and

[3] Theorem 17 states that "the capacity of a channel of band W perturbed by white thermal noise of power N when the average transmitter power is limited to P is given by

$$C = W \log \frac{P+N}{N}$$ (22, p. 67).

W is in cycles per second and takes the form of the reciprocal of some value of time. The power of a band of noise is equivalent to the variance of its amplitude distribution around its mean value.

a cross hair on an oscilloscope after the spot is suddenly displaced. Typically *S* is instructed to make an "exact" adjustment by the use of vision. The findings for this type of task indicate that the duration of quick corrective movements is a monotonically increasing function of amplitude. Studies of eye movements (6, 24) reveal a similar increase of movement duration with an increase in amplitude.

Although it has been known for many years that the accuracy of quick movements decreases with amplitude, the close relation of speed, amplitude, and accuracy often has been lost sight of. One of the earliest workers to emphasize the importance of their interrelation was Woodworth (25). Studying a variety of simple motor tasks performed with and without visual monitoring, he found that for quick, visually controlled movements variable error increases both with amplitude and with speed. Industrial engineers (19) also have consistently advocated that small-amplitude movements be used whenever possible in assembly work, on the grounds that the time required to complete a unit of work increases as a function of its amplitude as well as of the precision demanded by the task.

The concept of a fixed information-transmission capacity of the motor system not only accounts for such divergent results but also suggests a way of relating quantitatively the amplitude, duration, and variability of motor responses. The concept leads to the expectation that if repetitive movements of a fixed average amplitude are speeded up, then on the average each movement can provide less information, and movement variability will increase by a specified amount. Similarly, it suggests that if average movement amplitude should be increased then variability and/or average duration will also increase. The thesis can be tested experimentally in several ways.

An accepted method for estimating channel capacity in the case of a continuous signal requires a determination of the average power of the noise associated with the output signal. An experimental determination of the channel capacity of the human motor system by this method would require that the average amplitude and duration (or frequency) of movements be controlled by *E* and that the amplitude variability of *S*'s movements be measured by procedures similar to those used by Woodworth (25). However, the recording and measurement of successive movement amplitudes are laborious. Fortunately for the present purpose man has remarkably good ability to adjust his average rate of performance to meet the demands of a particular task, such as the requirement that his movements fall within some specified tolerance range. Thus it is possible to assess the tenability of the thesis of fixed information capacity by means of a test of the following specific hypothesis: *If the amplitude and tolerance limits of a task are controlled by* E, *and* S *is instructed to work at his maximum rate, then the average time per response will be directly proportional to the minimum average amount of information per response demanded by the particular conditions of amplitude and tolerance.*

In performing a given task *S* may actually use movements that are less variable, i.e., which generate more information, than the minimum demanded by the task. When only the limits of variability are specified, the minimum information condition is met if all amplitudes within the tolerance range are used equally often, i.e., when responses are distributed rectilinearly over the tolerance range, whereas the actual dispersion of response terminations is an approximate Gaussian distribution. The rate at which *S* makes successive responses in a given task, therefore, must be interpreted as the rate at which he can

generate, on the average, at least as much information per response as that demanded by the task. This restriction on the interpretation that can be given to results obtained with the present procedure is more than compensated for by the fact that the procedure makes it feasible to study a wide variety of task conditions with relatively large samples of *S*s and responses.

Experiment I: Reciprocal Tapping

In the first experiment two closely related reciprocal tapping tasks were studied. The *S*s were asked to tap two rectangular metal plates alternately with a stylus. Movement tolerance and amplitude were controlled by fixing the width of the plates and the distance between them. The task was accomplished primarily by movements of the lower arm.

Method

Apparatus and procedure.—A schematic drawing of the apparatus is shown in Fig. 1. The target plates were 6 in. long and of four different widths (2, 1, .5, and .25 in.). The widths of the plates (hereafter referred to as W_o) established the tolerance limits within which movements had to be terminated in amplitude. An additional error plate was mounted on either side of each target plate. The error plates were wide enough to catch all overshoots and undershoots. Six counters, activated by thyratron circuits through the plates and stylus, provided an automatic record of hits, overshoots, and undershoots for each set of three plates. After one of the counters in each group of three had recorded, none of the counters in that group would score again until one of the opposite set of plates had been touched. Thus the first plate touched on either side was always the one scored.

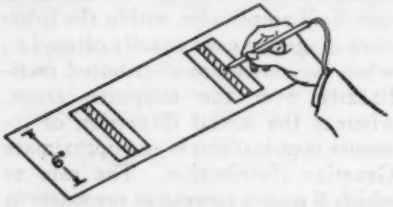


Fig. 1. Reciprocal tapping apparatus. The task was to hit the center plate in each group alternately without touching either side (error) plate.

Four center-to-center distances between the target plates were studied (2, 4, 8, and 16 in.) giving a total of 16 combinations of amplitude (A) and target size. (For $A = 2$ in. and $W_o = 2$ in. no undershoot errors were possible.)

Two metal-tipped styluses were used. One weighed 1 oz. and was about the size of a pencil. The other weighed 1 lb. and was slightly larger. The use of one or the other of these two styluses permitted the amount of work to be varied without otherwise changing the task.

The *S* sat with his right arm midway between the plates, which were located at a convenient height and angle. Each trial lasted 15 sec. and was followed by a 55-sec. rest period.

The following instructions were read to each *S*:

"Strike these two target plates alternately. Score as many hits as you can. If you hit either of the side plates an error will be recorded. You will be given a 2-sec. warning before a trial. Place your hand here and start tapping as soon as you hear the buzzer. *Emphasize accuracy rather than speed.* At the end of each trial I shall tell you if you have made any errors."

Subjects.—Sixteen right-handed college men served as *S*s.

Sequence of trials.—Each *S* was tested at each of the 16 conditions. A different random sequence of conditions was used for each *S* with the restriction that for the entire group of 16 *S*s each condition occur once on each trial. The experiment was replicated on the same day by retesting each *S* on each condition in a reversed sequence.

Each *S* participated in the experiment for two days. The lighter stylus was used on the first day, the heavier on the second day. All other procedures on the two days were the same. No attempt was made to equalize practice effects between the lighter and heavier styluses, because *E* was interested chiefly in the effects of amplitude and tolerance changes within each variation of the task.

Results

The time and error data for the two variations of the tapping task are included in Table 1. The data are expressed as the average time (t), in seconds, between taps (i.e., the time for each right-left or left-right movement) and as the percentage of errors (i.e., movements that were either too long or too short). Percentages are based on from 613 to 2,669 movements per condition.

TABLE 1

TASK CONDITIONS AND PERFORMANCE DATA FOR 16 VARIATIONS OF A RECIPROCAL TAPPING TASK

(N = the same 16 *Ss* at each condition)

Tolerance and Amplitude Conditions			1-oz. Stylus				1-lb. Stylus			
W_s	A	I_d	t	Errors (%)	I_p	Rank	t	Errors (%)	I_p	Rank
.25	2	4	.392	3.35	10.20	11	.406	3.80	9.85	7
.25	4	5	.484	3.41	10.33	9	.510	3.83	9.80	8
.25	8	6	.580	2.78	10.34	8	.649	4.04	9.24	13
.25	16	7	.731	3.65	9.58	14	.781	4.08	8.96	15
.50	2	3	.281	1.99	10.68	5	.281	0.88	10.68	4
.50	4	4	.372	2.72	10.75	3.5	.370	2.16	10.81	2
.50	8	5	.469	2.05	10.66	6	.485	2.32	10.31	6
.50	16	6	.595	2.73	10.08	12	.641	2.27	9.36	11
1.00	2	2	.212	0.44	9.43	15	.215	0.13	9.30	12
1.00	4	3	.260	1.09	11.54	1	.273	0.85	10.99	1
1.00	8	4	.357	2.38	11.20	2	.373	1.17	10.72	3
1.00	16	5	.481	1.30	10.40	7	.526	1.32	9.50	10
2.00	2	1	.180	0.00	5.56	16	.182	0.00	5.49	16
2.00	4	2	.203	0.08	9.85	13	.219	0.09	9.13	14
2.00	8	3	.279	0.87	10.75	3.5	.284	0.65	10.56	5
2.00	16	4	.388	0.65	10.31	10	.413	1.72	9.68	9

Note.—W_s is the width in inches of the target plate. A is the distance in inches between the centers of the two plates. t is the average time in seconds for a movement from one plate to the other. The performance index, I_p, is discussed in the text.

The average number of errors made in this task was small, only 1.2% with the lighter stylus and 1.3% with the heavier one. The low incidence of errors indicates that *Ss* successfully adjusted their rate of performance to meet changed task conditions.

The largest proportion of errors, 3.6% with the lighter stylus and 4.1% with the heavier stylus, was made at the most difficult task condition ($W_s = \frac{1}{4}$ in., $A = 16$ in.). At each amplitude step the percentage of errors for more exacting tolerances was consistently slightly larger than for less exacting tolerances; this indicated that *Ss* did not slow down their movements quite as much as necessary to effect the required accuracy at the more exacting conditions. Nor were the errors symmetrically distributed over the error plates. The *Ss* were uniformly more accurate in terminating flexor than extensor movements. With the lighter stylus over- and undershoot errors were equally frequent, but with the heavier stylus undershoots were about twice as frequent. However, in spite of such consistent trends, the relatively small incidence of errors justifies an analysis of the data in terms of the total number of movements made under the various task conditions.

Practice led to relatively small improvement in performance. The mean time per movement for all *Ss* and all task conditions decreased by 3% from the first to the second block of trials with the light stylus, and by 5% with the 1-lb. stylus. The total number of errors made by all *Ss* was almost exactly the same for the first as for the second block of trials.

For each category of W_s, movement time increased progressively as movement amplitude (A) increased. Likewise for each amplitude, movement time increased progressively as tolerance was decreased. The relations among the various conditions for the lighter and heavier styluses were similar.

Experiment II: Disc Transfer

The second task was to transfer plastic washers from one pin to another. Again, the amplitude and tolerance of the positioning movements were varied through a series of geometric steps, and the rate of performance was measured. This task differed from the previous one in that no errors were permitted. Some amount of finger activity was involved in grasping a washer by the edges before removing it from the pin.

Method

Apparatus and procedure.—A schematic drawing of the apparatus is shown in Fig. 2. Round plastic discs, ¼ in. thick and 1.5 in. in diameter, were drilled with four sizes of holes. Eight discs of uniform size were placed over a ½-in. pin. The task was to transfer the discs, one at a time, to another pin. The transfer of eight discs constituted a trial.

The difference between the diameter of the pins and the diameter of the center hole of the discs was varied in four geometric steps (⅛, ¼, ½, and 1 in.). Four amplitudes of movement (4, 8, 16, and 32 in.) were used, making a total of 16 conditions. The *S* touched a switch plate beside the empty pin as he began the task. This started a 1/100-sec. timer. The *S* touched the switch plate as soon as he had transferred the last disc. This stopped the timer. Discs were always moved from right to left.

Subjects.—Sixteen right-handed college men served as *S*s. They were different *S*s from those used in Exp. I.

Sequence of trials.—Each *S* worked at all 16 task conditions in a restricted randomized sequence corresponding to that used in the previous experiment. The entire procedure was then replicated on a second day using a different sequence of trials for each *S*.

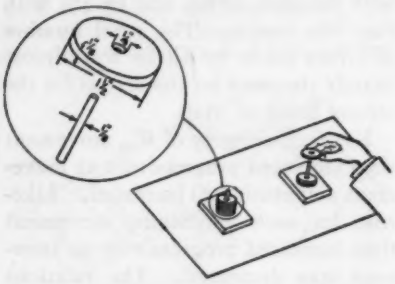


Fig. 2. Disc transfer apparatus. The task was to transfer eight washers one at a time from the right to the left pin. The inset gives the dimensions for the $W_s = \frac{1}{4}$ in. condition.

Results

The transfer of a disc involves a transport-empty plus a transport-loaded movement, and the times given are an average of these two aspects of each transfer cycle. There was a small (6%) decrease in time per movement on the second set of 16 trials. The average times per movement for each condition, averaged over all 16 *S*s and over both days, are given in Table 2 in the column labeled *t*. The variations in movement time with increase in movement amplitude and changes in tolerance correspond to those found in Exp. I.

Experiment III: Pin Transfer

The third task was to transfer pins from one set of holes to another. Grasping and releasing movements of the fingers were required in addition

TABLE 2

Task Conditions and Performance Data for 16 Variations of a Disc-Transfer Task

(N = the same 16 *S*s at each condition)

Tolerance and Amplitude Conditions			Performance		
W_s	A	I_d	t	I_p	Rank
.0625	4	7	.697	10.04	2.5
.0625	8	8	.771	10.38	1
.0625	16	9	.896	10.04	2.5
.0625	32	10	1.096	9.12	7
.125	4	6	.649	9.24	6
.125	8	7	.734	9.54	4
.125	16	8	.844	9.48	5
.125	32	9	1.028	8.75	10
.25	4	5	.607	8.24	12
.25	8	6	.672	8.93	9
.25	16	7	.771	9.08	8
.25	32	8	.975	8.20	13
.5	4	4	.535	7.47	16
.5	8	5	.623	8.02	14
.5	16	6	.724	8.29	11
.5	32	7	.902	7.76	15

Note.—W_s is the difference in inches between the diameter of the post and the diameter of the center hole of the disc. A is the center-to-center distance between the two posts. t is the average time in seconds for a movement between two posts.

to arm movements. No errors were permitted. The task was varied in four geometric steps of movement tolerance and in five geometric steps of amplitude. The time for completing a series of eight responses was recorded. The data for this task are somewhat more extensive, and somewhat more practice was given on this task than in the two preceding experiments.

Method

Apparatus and procedure.—The pin-transfer apparatus is shown in Fig. 3. Four sizes of pins ($\frac{1}{4}$, $\frac{1}{8}$, $\frac{1}{16}$, and $\frac{1}{32}$ in. in diameter) and five amplitudes of movement (1, 2, 4, 8, and 16 in.) were employed, thus making a total of 20 task conditions. Each set of pins was used with a set of holes whose diameter was twice that of the pins. The difference between the two diameters determined the tolerance (W_s) required in transferring a pin from one hole to another. The tips of the pins were tapered to one-half their diameter as shown in the inset of Fig. 3.

The *S* started and stopped a timer as in the preceding task, in order to record the over-all time for making eight transport-empty plus eight transport-loaded movements. A trial consisted of the transfer of eight pins from one side to the other followed after a short rest by the transfer of the same pins in the opposite direction.

The instructions were similar to those employed in the preceding experiments.

Subjects.—Twenty college students, ten men and ten women, served as *S*s. These were different *S*s from those serving in Exp. I and II.

Sequence of trials.—Each *S* served under all task conditions. The sequence of conditions was randomized, with the restriction that for the group of 20 *S*s each condition occur once on each trial and once for each *S*. The experiment was repeated three times with the same *S*s on three separate days. The sequence of trials for each *S* was different each day.

Results

There was some improvement in the rate of task performance from the first to the second day, but little further improvement on the third day. The over-all means for the three days were .613, .585, and .578 sec. per movement, respectively. The average movement times computed for the combined trials of the second and third replications of the experiment are given in Table 3 in the column labeled *t*. The relations among the data correspond closely to those found in the previous experiments. Movement times increased progressively as the tolerance requirement was made more stringent and as amplitude was increased.

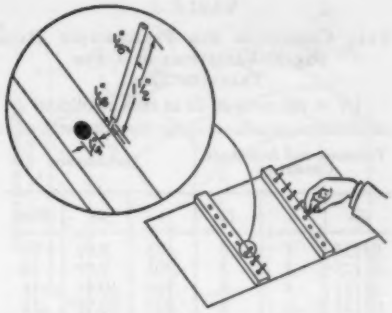


FIG. 3. Pin transfer apparatus. The task was to transfer eight pins one at a time from one set of holes to the other. The inset gives the dimensions of pins and holes for the $W_s = \frac{1}{4}$ in. condition.

INFORMATION ANALYSIS

The rate of performance of all the tasks studied increased uniformly as movement amplitude was decreased and as tolerance limits were extended. These results agree qualitatively with the predictions from the hypothesis.

In order to test the results against a quantitative prediction that the information output of the human motor system in any particular type of task is relatively constant over a range of amplitude and accuracy requirements, a difficulty index is needed that will specify the minimum information required on the average for controlling or organizing each movement.

The rational basis for this estimate of task difficulty is the maximum relative uncertainty that can be tolerated for a "correct" movement in a series having a specified average amplitude.[4] Thus the minimum organization re-

[4] Although only amplitude is considered here, the present rationale can be extended to the question of the rate of information involved in the selection of the direction, the force, or the duration of movements.

TABLE 3

TASK CONDITIONS AND PERFORMANCE DATA FOR 20 VARIATIONS OF A PIN-TRANSFER TASK

(N = the same 20 Ss at each condition)

Tolerance and Amplitude Conditions			Performance		
W_e	A	I_d	t	I_p	Rank
.03125	1	6	.673	8.92	20
.03125	2	7	.705	9.93	18
.03125	4	8	.766	10.44	14
.03125	8	9	.825	10.91	13
.03125	16	10	.959	10.43	15
.0625	1	5	.484	10.33	16
.0625	2	6	.518	11.58	11
.0625	4	7	.580	12.07	6
.0625	8	8	.644	12.42	2
.0625	16	9	.768	11.72	9
.125	1	4	.388	10.31	17
.125	2	5	.431	11.60	10
.125	4	6	.496	12.10	5
.125	8	7	.557	12.57	1
.125	16	8	.668	11.98	7
.25	1	3	.326	9.20	19
.25	2	4	.357	11.20	12
.25	4	5	.411	12.16	4
.25	8	6	.486	12.34	3
.25	16	7	.592	11.82	8

Note.—W_e is the difference in inches between the diameter of the pins and the diameter of the hole into which they were inserted. A is the center-to-center distance between the holes. t is the average time in seconds for a single movement.

quired of a particular movement is defined by the specification of one from among k possible categories of amplitude within which the movement is to terminate.

If we follow the above reasoning and conventional information notation, a binary index of difficulty (I_d) is defined as

$$I_d = -\log_2 \frac{W_e}{2A} \text{ bits/response}, \quad (1)$$

where W_e is the tolerance range in inches and A is the average amplitude of the particular class of movements. The choice of the denominator for this index is arbitrary since the range of possible amplitudes must be inferred. The use of $2A$ rather than A is indicated by both logical and practical considerations. Its use insures that the index will be greater than zero for all practical situations and has the effect of adding one bit ($-\log_2 \frac{1}{2}$) per response to the difficulty index. The use of $2A$ makes the index correspond rationally to the number of successive fractionations required to specify the tolerance range out of a total range extending from the point of initiation of a movement to a point equidistant on the opposite side of the target.

The index of difficulty (I_d) is given in Tables 1, 2, and 3 for each of the tasks. Since the index takes into account only the information necessary to specify the amplitude of a movement, it is insensitive to the weight of the stylus used in the first experiment and hence to the physical work required by the task. In its present form, the index is also insensitive to the information required for specifying the direction of a movement, or for specifying additional manipulatory acts such as the finger movements required to grasp and release objects.

Also shown in Tables 1, 2, and 3 is a binary index of performance (I_p), which expresses the results as a performance rate. For a task in which W_e and A are fixed for a series of movements, I_p is defined as

$$I_p = -\frac{1}{t} \log_2 \frac{W_e}{2A} \text{ bits/sec.}, \quad (2)$$

where t is the average time in seconds per movement. This index has some resemblance to one by Goldman (11, p. 157, Formula 29) for maximum information rate, although Goldman's criterion for determining the number of possible categories is based on a mean square criterion.

We want to test the hypothesis that movement time varies with task difficulty in such a way that I_p is constant over a wide range of movement amplitudes and tolerances. It is apparent from the data in Tables 1, 2, and 3 that even though the index is not precisely constant, the hypothesis is substantially confirmed.

For the eight best (out of the 16)

conditions of tapping with the light weight stylus, the rate of performance varied between 10.3 and 11.5 bits/sec., a range of only 1.2 bits. Performance fell off markedly only for the least exacting bits/sec., condition studied (A = 2 in., W_e = 2 in.).

Performance rate with the 1-lb. stylus also was relatively stable over a comparable range of amplitudes and tolerances. The rate at all but two conditions was reduced slightly by the added work, and the region of maximum performance was shifted toward smaller amplitudes of movement. The Pearsonian correlation between the 16 I_p values for the two variations in the tapping task was large however (r = .97).

For the 16 conditions of the disc-transfer task, on which the difficulty index ranged from 4 to 10 bits/response, the rate of performance varied from 7.5 to 10.4 bits/sec. Over the eight best conditions the range was only 1.3 bits/sec.

Performance on the pin-transfer task varied from 8.9 to 12.6 bits/sec. for the 20 conditions studied, but the range of variability was only 1.0 bit/sec. over the ten best conditions.

In all tasks it was found that movement amplitudes and tolerances could be varied within limits without much effect on performance rate, but that performance began to fall off outside these limits. This is readily apparent from an inspection of Fig. 4, which is a three-dimensional representation of the performance data for the pin-transfer task. Movements of 1- and 2-in. amplitude were consistently less efficient than movements of 4 to 8 in. In the pin-transfer task in particular, both the smallest amplitude (1 in.) and the smallest tolerance (1/32 in.) resulted in relatively low rates of performance. In most cases performance rate also fell off at the largest amplitudes and tolerances. Throughout all the tasks, movement amplitudes of 4 to 8 in. were more consistently associated with good performance than was any particular tolerance or any particular difficulty index.

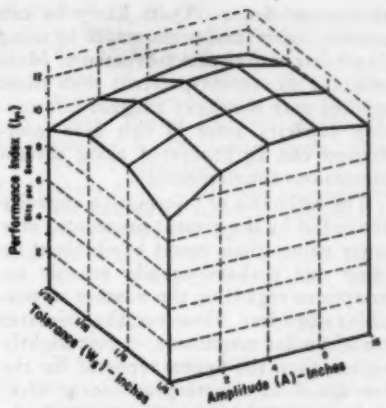


FIG. 4. Three-dimensional representation of performance rate (I_p) in bits per second for the pin-transfer task as a function of amplitude and tolerance requirements

DISCUSSION

Although the four tasks thus far analyzed are a small sample of perceptual-motor activities, the results are sufficiently uniform to indicate that the hypothesized relation between speed, amplitude, and tolerance may be a general one. It is not surprising to find that performance, measured in information units, is relatively constant over a central range of amplitude and accuracy conditions and falls off outside these limits. Such a relation holds for most perceptual and motor activities and has recently been found (16) in a study of the optimum relation between the amplitude of lever movement and observed display movement.

The absolute level of capacity of the motor system probably varies considerably for different movements, limbs, and muscle groups. For example, the arm, which in the present experiments was asked to generate information only in respect to successive movement amplitudes, may have a lower information capacity than the hand. Using all of his fingers singly and in combinations, S can produce one bit of information per re-

sponse per finger. Quite likely he can sustain higher level performance by using finger rather than arm movements. More complex movement patterns than those studied may also have a higher information capacity since in this case information can be generated along several dimensions simultaneously.

The estimates of information capacity arrived at by the present procedures give lower values than would be obtained by other and perhaps equally tenable assumptions regarding the number of possible categories. However, the estimates are of similar magnitude, in fact slightly higher, than the figures reported for the average *S* in a perceptual-motor task, such as pressing keys in response to flashing lights (18), in which responses to a series of discrete stimuli have been studied.

The present index of difficulty can be applied to a wide range of motor tasks. For example the three stick and rudder control movements necessary to make a single match on the Complex Coordination Test (Model E) require one bit of information to specify each of the three directions of movement and 3.9 bits to specify each of the three amplitudes (a total of 14.7 bits); this assumes that *S* starts his adjustments with the controls in the neutral position.

Ellson (7) has raised the question of the linearity of motor responses, using the superposition theorem as the criterion of linearity. This theorem demands that movement duration remain constant as amplitude increases. We have noted that many recent data violate this requirement. However, the present results indicate that over the range in which performance is optimum, movements of differing amplitude but of equal difficulty in terms of information tend to be of approximately equal duration. Man differs from a linear electronic or mechanical system, however, in that he sometimes varies his standard of relative precision for a movement at the same time as he varies its amplitude; unless relative precision remains constant, movement time will vary with amplitude.

The basis for the uniformity of motor performance perhaps is to be found in the time taken for central organizing processes. Fenn has pointed out that in reciprocal movements made at maximum rate and without any visual control the muscles work at a level far below their physical capacity. Since maximum rate of performance in this case is not limited by the muscles, he has proposed that the maximum rate is limited by ". . . the speed with which excitation and inhibition can be made to alternate in the central nervous system without at the same time losing precise control of the magnitude of the force" (8, p. 174). This concept of central organizing time is entirely consistent with an information point of view. The finding that relatively small differences in performance result from the change in stylus weight, and the validity of predictions of performance rate from the index of task difficulty lend support to the basic thesis of this paper, that it is the degree of control required over the organization of a response, i.e., the amount of information required to specify a response, which is the major factor limiting the rate of motor performance.

Summary

The present paper attempts to relate the traditional Weber function (variability of a response as a function of its amplitude) to the parallel phenomena of variability as a function of response duration, using certain concepts of information theory.

An index of the difficulty of a movement is proposed on the assumption that the average amplitude, the average duration, and the amplitude variability of successive movements are related in a manner suggested by information theory. The basic rationale is that the minimum amount of information required to produce a movement having a particular average amplitude plus or minus a specified tolerance (variable error) is proportional to the logarithm of the ratio of the tolerance to the possible amplitude range. The specification of the possible amplitude range is arbitrary and has been set at twice the average amplitude. The average rate of information generated by a series of movements is the average information per movement divided by the time per movement.

Three experiments are reported which were designed to test the following hypothesis: If the amplitude and tolerance limits of a task are fixed and the *S* is instructed to work at his maximum rate, then the average duration of responses will be directly proportional to the minimum average amount of information per response (i.e., the

degree of behavior organization) demanded by the task conditions. The conditions studied covered the range from 1 to 10 bits/response.

The results indicate that rate of performance in a given type of task is approximately constant over a considerable range of movement amplitudes and tolerance limits, but falls off outside this optimum range. The level of optimum performance was found to vary slightly among the three tasks in the range between about 10 to 12 bits/sec. The consistency of these results supports the basic thesis that the performance capacity of the human motor system plus its associated visual and proprioceptive feedback mechanisms, when measured in information units, is relatively constant over a considerable range of task conditions. This thesis offers a plausible way of accounting for what otherwise appear to be conflicting data on the durations of different types of movements.

The author feels that the fixed information-handling capacity of the motor system probably reflects a fixed capacity of central mechanisms for monitoring the results of the ongoing motor activity while at the same time maintaining the necessary degree of organization with respect to the magnitude and timing of successive movements.

References

1. Brown, J. S., & Slater-Hammel, A. T. Discrete movements in the horizontal plane as a function of their length and direction. *J. exp. Psychol.*, 1949, **39**, 84–95.
2. Bryan, W. L. On the development of voluntary motor ability. *Amer. J. Psychol.*, 1892, **5**, 125–204.
3. Craig, D. R. *Effect of amplitude range on duration of responses to step function displacements.* Dayton, O.: U. S. Air Force, Air Materiel Command, Wright-Patterson Air Force Base, 1949. (AF Tech. Rep. No. 5913.)
4. Cronbach, L. J. *A consideration of information theory and utility theory as tools for psychometric problems.* Urbana, Ill.: Bureau of Research and Service, Univ. of Illinois, 1953.
5. Crossman, E. R. F. W. Entropy and choice time: the effect of frequency unbalance on choice-responses. *Quart. J. exp. Psychol.*, 1953, **5**, 41–51.
6. Dodge, R., & Cline, T. S. The angle velocity of eye movements. *Psychol. Rev.*, 1901, **8**, 145–157.
7. Ellson, D. G. The application of operational analysis to human motor behavior. *Psychol. Rev.*, 1949, **56**, 9–17.
8. Fenn, W. O. Mechanics of muscular contraction in man. *J. appl. Physics*, 1938, **9**, 165–177.
9. Freeman, F. N. Experimental analysis of the writing movement. *Psychol. Monogr.*, 1914, **17**, No. 4 (Whole No. 75), 1–46.
10. Garner, W. R., & Hake, H. W. The amount of information in absolute judgments. *Psychol. Rev.*, 1951, **58**, 446–459.
11. Goldman, S. *Information theory.* New York: Prentice-Hall, 1953.
12. Hartson, L. D. Contrasting approaches to the analysis of skilled movements. *J. gen. Psychol.*, 1939, **20**, 263–293.
13. Hick, W. E. On the rate of gain of information. *Quart. J. exp. Psychol.*, 1952, **4**, 11–26.
14. Hick, W. E., & Bates, J. A. V. *The human operator of control mechanisms.* London: Ministry of Supply, 1950.
15. Howes, D. H., & Solomon, R. L. Visual duration threshold as a function of word-probability. *J. exp. Psychol.*, 1951, **41**, 401–410.
16. Jenkins, W. L., & Olson, M. W. The use of levers in making settings on a linear scale. *J. appl. Psychol.*, 1952, **36**, 269–271.
17. Klemmer, E. T., & Frick, F. C. Assimilation of information from dot and matrix patterns. *J. exp. Psychol.*, 1953, **45**, 15–19.
18. Klemmer, E. T., & Muller, P. F., Jr. The rate of handling information. Key pressing responses to light patterns. *USAF Hum. Fact. Operat. Res. Lab.*, 1953, Rep. No. 34.
19. Maynard, H. B., Stegemerten, G. J., & Schwab, J. L. *Methods—time measurement.* New York: McGraw-Hill, 1948.
20. Miller, G. A. What is information measurement? *Amer. Psychologist*, 1953, **8**, 3–11.
21. Searle, L. V., & Taylor, F. V. Studies in tracking behavior. I. Rate and time characteristics of simple corrective movements. *J. exp. Psychol.*, 1948, **38**, 615–631.
22. Shannon, C., & Weaver, W. *The mathematical theory of communication.* Urbana: Univ. of Illinois Press, 1949.
23. Stetson, R. H., & McDill, J. A. Mechanism of the different types of movement. *Psychol. Monogr.*, 1923, **32**, No. 3 (Whole No. 145), 18–40.
24. Westheimer, G. The response of the oculomotor system to visual stimuli in the horizontal plane. Unpublished doctor's dissertation, Ohio State Univ., 1953.
25. Woodworth, R. S. The accuracy of voluntary movement. *Psychol. Monogr.*, 1899, **3**, No. 2 (Whole No. 13).

(Received for early publication March 13, 1954)

Journal of Experimental Psychology
Vol. 47, No. 6, 1954

THE RELATION OF DIFFERENT LEVELS AND KINDS OF MOTIVATION TO VARIABILITY OF BEHAVIOR [1]

RUSSELL L. DE VALOIS

University of Michigan

For some time a great amount of interest has centered on the problem of the extremely invariable behavior (variously termed stereotypy, fixation, or rigidity) exhibited by animals in certain situations. It has been reported (**4, 5, 7, 15, 16**) that rats shocked at or before a choice point in a maze tend to fixate on a particular response, exhibiting virtually no variability in choice of alley from trial to trial. It has also been found (**7, 16**) that this invariable behavior tends to persist after the shock has been discontinued in the maze, and even if *Ss* are shocked in the goal for making the fixated response.

In a series of related experiments, Maier and his colleagues have studied fixated behavior exhibited by rats in the Lashley jumping apparatus when they were given various types of problems to solve.

An important feature of these experiments was that if the *Ss* hesitated, they were forced to jump from the stand, by the administration of a blast of air from the rear. It was found (**6, 9, 10, 11**) that when *Ss* were presented with insoluble problems in which, no matter what choice they made, they would find the door locked half the time, a considerable number of them fixated, i.e., made consistent, invariable choices from trial to trial. This behavior continued even after the problems were made soluble and more adaptive responses were possible. Under these circumstances *Ss* had to be forced to jump by the administration of the air blast. It was also reported (**6**) that some *Ss* fixate when given an air blast even though there is no problem (both doors always unlocked), and that some fixate in an insoluble problem even when not forced, with the air blast, to jump.

Maier has interpreted these experiments by postulating two different behavior mechanisms: motivation and frustration. In motivated behavior rewards and punishments strengthen and weaken responses respectively; the behavior is goal oriented. In frustrated behavior punishment does not weaken responses but, if anything, strengthens them. Thus, frustrated *Ss* do not change their responses when they are punished for making them, or when more adaptive responses are available. McClelland (**8**), objecting to frustration theory, has pointed out that the frustrated *Ss* of Maier's experiments can be considered as *Ss* operating under avoidance motivation, i.e., escape from the air blast, or from the punishing situation of jumping to locked doors in an insoluble problem. He has hypothesized that approach and avoidance motivations lead to different amounts of variability, avoidance-motivated behavior being much more rigid or less variable than approach-motivated behavior.

The theory being tested in this experiment is that amount of variability is a function of the amount of motivation rather than of the type of motivation. There is good evidence (**2**) that more intense motivation can be provided in escape or avoidance situa-

[1] This article is based on a portion of a dissertation submitted in February 1952 to the graduate school of the University of Michigan in partial fulfillment of the requirements for the Ph.D. degree. The writer is indebted to the members of the committee and to Dr. Edward L. Walker, chairman of the committee, who gave generous aid and invaluable criticism.

tions than in approach situations. Klee (6) has found that some rats would starve to death rather than jump in an insoluble problem, but that they could be made to jump with an air blast. Thus, the experiments showing great rigidity of behavior in shock- or air blast-motivated *Ss* would be interpreted as being the result of the more intense motivation of these *Ss* than of the hunger-motivated ones.

The central problem of this experiment is the relationship of amount of variability to strength of motivation. This was studied for an approach motive (thirst) and for an avoidance motive (escape from shock) to determine the extent to which the relationship holds for two different types of motivation. It is also of interest to determine whether the range of variability found under the approach and avoidance motivations shows any overlap. Furthermore, the effect on variability of changing the motivation of *Ss* was studied to determine the applicability of theories of secondary motivation to this problem.

Method

Subjects

A total of 82 experimentally naive male albino rats obtained from the Dixboro Farm were used as *Ss*. The *S* were 70 to 80 days of age at the start of the experiment and had about 30 days of handling prior to that time.

Apparatus

Diagrams of the whole maze and of details of a single unit are shown in Fig. 1. At each of the choice points *S* could go either right, left, or straight ahead (if the middle alley was open). The maze was enclosed by wooden sides, a glass top, and a grid floor. Galvanized steel wires running across the alleys for the entire maze, including the goal box, served as the floor and could be electrified to serve as a shock grid. The walls of the maze stood on a table, and the grid wires ran through the walls .5 in. above the table top. The over-all length of the maze was approximately 21 ft., each of the connecting alleys was 18 in., and the distance across each diamond was 28 in.

Sliding masonite doors (D), which *E* could manipulate by means of strings, were placed as shown in Fig. 1. These were used to prevent backtracking, and to stop *S* briefly before it entered each diamond. In addition, semipermanent glass doors (G) were placed .5 in. from either end of each middle alley. These were used to block passage into the middle alleys of the diamonds during certain parts of the experiment.

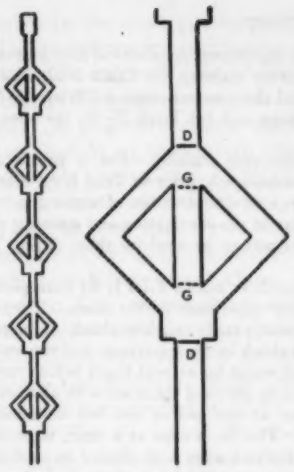


FIG. 1. Diagram of the complete maze, with details of a single unit

The starting alley consisted of a regular connecting alley leading to the first diamond. The goal box was 8 in. × 13½ in. The roof was of glass for the shock *Ss* and of wire mesh for the water *Ss*. Water was provided for the latter by means of a glass tube projecting into the goal box from a bottle on top of the mesh roof. Both the glass and wire mesh tops were hinged at the maze end and could be tilted to remove *Ss* from the goal box.

The walls of the maze were made of pine wood and, except as otherwise noted, were left unfinished. Some sections were painted a flat black. The painted sections were the starting alley, all of the first diamond, the connecting alley to the second diamond, all of the second diamond, the left outside alley of the third diamond, the right outside alley of the fourth diamond, and the connecting alley from the fifth diamond to the goal box.

The shocking circuit which was connected to the grid floor of the maze was an electronic circuit designed to give a constant current output despite changes in resistance across the leads. The voltage would change automatically whenever there was a change in the resistance across the grid, in order to maintain a constant current. The current produced was a 120-cycle pulsating direct current.

Procedure

The experiment consisted of four periods: (*a*) preliminary training, (*b*) Trials 1–12, which will be called the constant stage, (*c*) Trials 13–24, the shift stage, and (*d*) Trials 25–36, the open maze stage.

Preliminary training.—For a period of six days immediately prior to Trial 1, *S*s that were to be run on a motivation of water deprivation were put on the deprivation and watering schedule that was to be used for them during Trials 1–12.

Three days before Trial 1, *S*s were given exploratory experience in the maze. They were run unmotivated, i.e., the shock *S*s were not given a shock in the apparatus, and the water *S*s had had water for several hours before running. The middle alleys of the maze were all open, but the door at the end of the last diamond was closed. The *S*s, a cage at a time, were placed at the starting alley and allowed to explore the maze for a period of 3 hr., after which they were removed by *E*.

Two days before Trial 1, all *S*s were given preliminary training in running to the goal, under the motivating conditions that were to obtain in Trials 1–12. Each *S* was given two trials during which it ran from the end of the last diamond through the connecting alley into the goal box. Water was available in the goal box for the water *S*s. They were left in the box until they discovered the water and drank, and then were removed to a cage at the end of the maze. Each *S* in a cage was run through this procedure in turn, and then each was run again. In the case of the shock *S*s, *E* put them into the connecting alley, closed the door, and turned on the shock. Upon reaching the unelectrified goal box, *S*s were allowed to explore it briefly before being removed to the terminal cage.

On the final day of the preliminary training, *S*s were run through the connecting alley to the goal in exactly the same manner as in the previous day, except that they were given only one trial.

Experimental trials.—The *S*s were given one trial per day under one of four different motivating conditions. The motivating conditions employed were 6-hr. water deprivation, 22-hr. water deprivation, light shock (40 microamp.), and strong shock (500 microamp.). The water *S*s were rewarded with unlimited amounts of water in the goal box. The shock *S*s ran through the maze with the entire floor electrified, and were rewarded by the cessation of the shock when they reached the unelectrified goal box.

During the first 24 trials, the middle alleys were all closed by the use of the glass doors. The *S*s could, therefore, go either right or left at each choice point, but not straight ahead. During the first 12 of these trials there were four groups of *S*s, the groups being differentiated according to the motivating conditions. On Trial 13 each group was divided in half, one half being continued for the rest of the experiment under the same motivation as in the first 12 trials, and the other half being shifted to another level of motivation for the rest of the experiment. Thus, there were four groups of *S*s during Trials 1–12, and eight groups in Trials 13–36. The *S*s whose motivation was shifted after Trial 12 were shifted from 6- to 22-hr. water deprivation, from 22- to 6-hr. deprivation, from light to strong shock, and from strong to light shock. We have designated these various groups as 6–6, 6–22, 22–22, 22–6, L–L, L–S, S–S, S–L.

On Trial 25 the middle alleys through the maze were opened by removal of the glass doors. During Trials 25–36, therefore, *S*s could continue along their previous paths or considerably shorten their run to the goal by taking one or more segments of the middle straightaway.

Results

The experiment consisted, in a sense, of two subexperiments using the same *S*s. One of these experiments was the closed middle alley situation (Trials 1–24, the constant and shift stages), and the second, the open maze stage (Trials 25–36). The results will be presented accordingly.

Closed middle alley situation.—Although several methods of measuring behavior variability are available with our data, the measure chosen is the number of choice-point switches. Each choice point is considered as a unit, and the behavior on successive days at this juncture is studied. The *S* can, at the first of the five choice points, go either to the same side as it did the previous day or to the opposite side. If it goes to the opposite side, it is credited with one choice-point switch. The other four choice points are considered similarly, and *S* is given a score for the day indicating the number of choice points at which it switched to the opposite side on this day with respect to the previous day. Possible scores, then, are 0 through 5 for each *S* each day. A high choice-point switch score indicates high inter-trial variability.

TABLE 1

CHOICE-POINT SWITCHES IN TRIALS 1–24, CONSTANT MOTIVATION GROUPS

Group	*N*	Mean	*SD*	*t*
6–6	10	46.60	12.86	2.43* (6, 22)
22–22	11	31.36	14.26	4.49** (22, S)
L–L	10	31.20	12.13	2.61* (L, 6)
S–S	10	8.00	6.64	5.02** (S, L)

* Significant beyond the .05 level.
** Significant beyond the .01 level.

The results for those *Ss* on constant motivation throughout the experiment are shown in Table 1. All the differences between groups are significant with the exception of that between the strong thirst and the light shock *Ss*. The less highly motivated *Ss* in both the thirst and shock groups are significantly more variable than the more highly motivated *Ss*. Since there is good reason to believe that the range of motivation obtainable with shock is higher than with water deprivation, a continuum of strength of motivation might be assumed going from 6-hr. thirst through 22-hr. thirst and light shock to strong shock. Then these results indicate that the more intense the motivation the less the behavioral variability. There is nothing in the data which would require any difference between thirst and shock motivation other than the quantitative difference between their strengths.

On Trial 13 half of the *Ss* in each group were shifted to a different level of motivation. The point that is of interest here is the effect upon variability of shifting the motivational level. Toward this end, comparisons are made between the change in performance, from the constant to the shift stage, of *Ss* kept on constant motivation and those whose motivation is changed. In Table 2 the mean changes in the number of choice-point switches from the constant to the shift stage of the various groups of *Ss* are presented. For each *S* the number of switches in the constant stage (Trials 1–12) was subtracted from the number in the shift stage (Trials 13–24), and the means for each of the groups computed. A negative score, therefore, means that *Ss* in that group were less variable in the second 12 trials than they were in the first 12 trials.

It can be seen that those *Ss* shifted from 6 to 22 hr. of thirst changed to a lower variability than those maintained on 6-hr. deprivation although the difference was not significant, and those shifted from 22 to 6 hr. of thirst became significantly more variable than those maintained on 22-hr. deprivation. The *Ss* whose shock-escape motivation was increased shifted in the same direction as those whose thirst motivation was increased, but this shift was not significant. Decreasing the amount of shock that the strong shock *Ss* were given had no significant effect on their variability.

Open maze stage.—This part of the experiment involves essentially the sort of situation that Maier and his colleagues have used in studying a different type of variability: fixation, or inability to modify behavior when a more adaptive response is possible. Before Trial 25 the middle alleys through the maze were opened, and *Ss* could now go straight to the goal

TABLE 2

CHANGES FROM THE CONSTANT STAGE TO THE SHIFT STAGE IN THE MEAN NUMBER OF CHOICE-POINT SWITCHES

Group	*N*	Mean	*SD*	*t*
6–6	10	.80	5.71	1.47
6–22	10	−3.90	7.70	
22–22	11	−3.73	5.60	3.32*
22–6	11	8.27	9.96	
L–L	10	−3.00	5.37	1.08
L–S	9	−6.78	8.80	
S–S	10	−4.40	4.84	.22
S–L	11	−4.64	3.02	

* Significant beyond the .01 level.

TABLE 3

MIDDLE ALLEY ENTRANCES IN TRIALS 25–30

Group	Mean	*SD*	*t*
6–6	13.20	6.49	2.58* (6–6, 22–22)
22–6	12.82	7.21	2.39* (22–6, 22–22)
22–22	5.45	6.71	2.45* (22–22, S–S)
6–22	1.50	2.70	5.09** (6–22, 6–6)
L–L	13.70	7.89	2.46* (L–L, 22–22)
S–L	1.64	2.36	2.08* (S–L, S–S)
S–S	0.00	.00	5.50** (S–S, L–L)
L–S	0.33	.66	4.81** (L–S, L–L)

* Significant beyond the .05 level.
** Significant beyond the .01 level.

through the middle alleys. In order to assess variability of behavior in terms of the absence of "fixation," the number of entries into the middle alleys was determined for each *S*.

In Table 3 the mean number of middle alley entries for the eight separate groups are given for Trials 25–30. It is to be noted that the second symbol in the designation of each group indicates the motivating condition during these trials. These trials rather than the whole 12-trial period were analyzed because it was felt that the number of middle alley entries during the early period is a measure of the extent to which *S*s are breaking away from their previous habit, whereas the number of entries later in the period would indicate, rather, the extent to which *S*s had by then stereotyped on this response. However, relationships among the groups are the same whether Trials 25–30 or Trials 25–36 are analyzed.

It can be seen that the results from the two different parts of the experiment, studying two different types of variability, are in agreement with the exception of the behavior of the light shock *S*s. In the case of *S*s on constant motivation throughout the experiment, the less highly motivated *S*s within each type of motivation are more variable than the more highly motivated ones. The thirst *S*s whose motivation had been increased or decreased were less or more variable, respectively, than the corresponding *S*s on constant motivation. The constant light shock *S*s in this part of the experiment, however, were significantly more variable than the strong thirst *S*s, whereas previously there was not a significant difference between the groups. The shock *S*s whose motivation had been increased were significantly less variable than the corresponding constant motivation group, whereas this difference in the first part of the experiment, while in the same direction, was not significant.

DISCUSSION

The major purpose of this investigation was to ascertain whether degree of variability of behavior within a kind of motivation could be related to the level or strength of a motive. Secondly, it was desired to determine whether the same relation between degree of variability and level of motivation pertained in the case of a typical approach motive and a typical avoidance motive. Thirdly, it was hoped that by appropriate choices of motivational level, the range of variability within an approach motive could be made to overlap with the range of variability within an avoidance motive. If this could be done, it was felt that the necessity for a dichotomous relationship between the effects of approach and avoidance drives, and possibly between the mechanisms of motivation and frustration, as Maier defines them, could be reduced. Finally, by shifting the motivational level within each kind, it was hoped that additional information could be secured relative to the persistence of stereotypy as a function of level of motivation and perhaps upon the applicability of theories of secondary motivation to these problems.

The evidence obtained allows us to come to definite conclusions with regard to the major problem of the experiment. The weak thirst *S*s were significantly more variable than the strong thirst *S*s in both parts of the experiment. The weak shock *S*s were also significantly more variable than the strong shock *S*s. High

levels of motivation, then, within each kind of motivation resulted in less variability than low levels of motivation.

Positive evidence was obtained with regard to the second problem of the experiment, that of whether the same relationship of amount of variability to motivational level held within each of the kinds of motivation used. As had been indicated above, increased motivation led to decreased amount of variability in the case of both the thirst *S*s and the shock escape *S*s.

With regard to the problem of the overlap of the range of variability of approach- and avoidance-motivated *S*s, no significant difference was found between the variability of the strong thirst and the light shock *S*s in the first part of the experiment, and the light shock *S*s were significantly more variable than the strong thirst *S*s in the second part of the experiment. There is thus good evidence against a dichotomy of variability between approach- and avoidance-motivated *S*s.

Another problem concerns what effect changing the level of motivation has upon the amount of variability exhibited. If variability of behavior were an inverse function of the present motivation of *S*, one would expect that increasing the motivation would lead to a decreased variability. This is what was found for both thirst and shock *S*s, except that the L–S group was not significantly less variable than the L–L group in the first part of the experiment, although the difference was in the predicted direction. In the case of decreasing the amount of motivation one runs into complications, which theories of secondary motivation have been developed to explain (13, 14). It would be predicted from secondary motivation theory that intensely motivated *S*s would acquire secondary motivation that would continue to operate, until gradually extinguished, as a strong drive after the original intense motivation was discontinued or decreased. If this theory were extended by postulating that the intensity of the secondary motivation acquired is a function of the intensity of the original motivation, the results of this experiment regarding those groups whose motivation was decreased would be explainable. The strong shock *S*s, having a very intense motivation, would increase in variability less rapidly when their primary motivation was decreased than would the less intensely motivated strong thirst group. This corresponds to the results obtained. There is also other supporting evidence (1, 12).

In an experiment somewhat related to certain parts of this experiment, Elliott (3) reported that (*a*) increased hunger leads to decreased variability, and (*b*) *S*s that have become invariable as a result of increased hunger do not return to their previous level of variability when their hunger is reduced. The second result does not agree with the results of this experiment. We believe that the contradictory results are the result of the technique Elliott used to motivate his animals. His *S*s were fed 5/6 of their daily ration for a period of time (low motivation level), and then for ten days were fed 1/3 of their ration (high motivation level), and finally were returned for ten days to the low motivation level. It would appear that *S*s would be hungrier, because of the accumulated food deficit, at the low motivation level the second time than they were during the original runs at this level, and thus would be less variable in behavior. If this is granted, there is no contradiction between his findings and those reported herein, and his second conclusion would not be substantiated.

It is believed that the theory, which states that variability of behavior is a function of the amount of motivation, can adequately explain the results of experiments which have shown that *S*s exhibit fixated behavior when shocked in mazes, or given insoluble problems and/or air blasts in a jumping apparatus. These various situations are seen as providing very intense motivation for *S*s, and therefore producing less variable behavior than the usual thirst or hunger motivation. The results obtained in this experiment, particularly the different levels of variability exhibited by groups operating under the same type of motivation and the overlap in the amount of variability exhibited by the shock and

the thirst *Ss*, do not agree with theories which postulate a dichotomy of variable and fixated behavior; these two types of behavior are produced by different mechanisms. It appears to us that since there is no evidence here for a dichotomy, and it is possible to explain the results of other experiments in terms of a single variable—strength of motivation—it would be more parsimonious not to introduce two different mechanisms of behavior—motivation and frustration or approach and avoidance motivation—but rather to treat the experimental data in terms of different amounts of motivation.

Summary

Groups of rats operating under four different motivating conditions, 6-hr. thirst, 22-hr. thirst, light shock, and strong shock, were run through a maze in which there were a number of equally long routes to the goal. Half of the *Ss* within each group were shifted, half way through the experiment, to the other motivational level within their type. Later in the experiment, in a study of a different type of variability, a new shorter path to the goal was made available.

1. The low motivation *Ss* within each type of motivation were more variable than the high motivation *Ss*.

2. There was an overlap in the amount of variability exhibited by the approach and the avoidance groups.

3. Increasing the strength of motivation leads to a decreased amount of variability.

4. Decreasing the strength of motivation leads to an increased amount of variability in the case of *Ss* originally run on moderately strong motivation (strong thirst), but not significantly in the case of *Ss* originally run on extremely high motivation (strong shock).

These results are held to be consonant with a theory that amount of variability is an inverse function of the amount of motivation. The effect on variability of changing the level of motivation is explicable in terms of this relationship and secondary motivation theory. No reason is seen for postulating a dichotomy of variable and fixated behavior, with different mechanisms for the two types of behavior.

References

1. Anderson, E. E. The externalization of drive. II. The effect of satiation and removal of reward at different stages in the learning process of the rat. *J. genet. Psychol.*, 1941, **59**, 359–376.

2. Brown, J. S. Gradients of approach and avoidance responses and their relation to level of motivation. *J. comp. physiol. Psychol.*, 1948, **41**, 450–465.

3. Elliott, M. H. The effect of hunger on variability of performance. *Amer. J. Psychol.*, 1934, **46**, 107–112.

4. Everall, E. E. Perseveration in the rat. *J. comp. Psychol.*, 1935, **19**, 343–369.

5. Hamilton, J. A., & Krechevsky, I. Studies in the effect of shock upon behavior plasticity in the rat. *J. comp. Psychol.*, 1933, **16**, 237–253.

6. Klee, J. B. The relation of frustration and motivation to the production of abnormal fixations in the rat. *Psychol. Monogr.*, 1944, **56**, No. 4 (Whole No. 257).

7. Kleemeier, R. W. Fixation and regression in the rat. *Psychol. Monogr.*, 1942, **54**, No. 4 (Whole No. 246).

8. McClelland, D. C. *Personality*. New York: Sloan, 1951.

9. Maier, N. R. F. *Frustration: the study of behavior without a goal*. New York: McGraw-Hill, 1949.

10. Maier, N. R. F., Glaser, N. M., & Klee, J. B. Studies of abnormal behavior in the rat. III. The development of behavior fixations through frustration. *J. exp. Psychol.*, 1940, **26**, 521–546.

11. Maier, N. R. F., & Klee, J. B. Studies of abnormal behavior in the rat. XII. The pattern of punishment and its relation to abnormal fixations. *J. exp. Psychol.*, 1943, **32**, 377–398.

12. Miller, N. E. Experiments on the strength of acquired drives based on hunger. *Amer. Psychologist*, 1947, **2**, 303. (Abstract)

13. Miller, N. E. Studies of fear as an acquirable drive. I. Fear as motivating and fear-reduction as reinforcement in the learning of new responses. *J. exp. Psychol.*, 1948, **38**, 89–101.

14. Mowrer, O. H., & Lamoreaux, R. R. Avoidance conditioning and signal duration—a study of secondary motivation and reward. *Psychol. Monogr.*, 1942, **54**, No. 5 (Whole No. 247).

15. Muenzinger, K. F., & Wood, A. Motivation in learning. IV. The function of punishment as determined by its temporal relation to the act of choice in the visual discrimination habit. *J. comp. Psychol.*, 1935, **20**, 95–106.

16. O'Kelly, L. I. An experimental study of regression. I. Behavioral characteristics of the regressive response. *J. comp. Psychol.*, 1940, **30**, 41–53.

(Received August 21, 1953)

Journal of Experimental Psychology
Vol. 47, No. 6, 1954

AN EXPERIMENTAL DISTINCTION BETWEEN PERCEPTUAL PROCESS AND VERBAL RESPONSE[1]

ULRIC NEISSER

Massachusetts Institute of Technology

Data from perceptual experiments are necessarily behavioral. We may be interested in what the subject sees, but we can only record what he says or does. This situation gives rise to a number of theoretical problems. Some of these problems concern the very subject matter of the field of perception. Thus, while Graham (1) suggests that perceptual experiments are essentially concerned with certain types of responses or behavior functions, Osgood (5, p. 194) discusses "awareness, as . . . indexed by verbal or other modes of response."

Broad differences in approach are not the only consequences of the mediated character of perceptual data. More concrete issues are affected. For example, we know that duration thresholds for tachistoscopically presented words are a function of language frequencies (2); but we do not know whether the more frequent words are more readily *seen* or *said*. A related problem concerns the duration thresholds of words with emotional connotations (4).

The present experiment is intended to demonstrate that the effect of preparatory set in tachistoscopic experiments is to facilitate seeing rather than saying the words presented. That is, a set does not raise the probabilities of the critical responses generally, but only under restricted stimulus conditions; namely, when the expected words are shown.

If the demonstration is valid, it suggests a corresponding interpretation of the other tachistoscopic studies. For example, the effect of language frequency may then also be supposed to be at the stimulus end of the process. On the other hand, these data are not presented as evidence relevant to the fundamental problem of perception and response. Indeed, even a discussion of the possibility of such evidence is beyond the scope of this paper. At best, demonstrations like the one presented here may be useful in some future clarification of the issue.

The demonstration depends upon the use of homonyms: words that are spelled differently yet are pronounced alike. A set is induced for one member of a homonym pair, and the consequent decrease in threshold for that member is compared with the decrease for the second member. The same verbal response (within the limits of experimental measurement) is employed in reporting either word of the pair. Therefore, recognition of the second member will be facilitated if the set has operated directly on response probabilities, but will not be facilitated if the set affects perceptual processes.

METHOD

Subjects.—The *S*s were 12 male undergraduates at the Massachusetts Institute of Tech-

[1] This research was supported in part by the United States Air Force under Contract No. AF18(600)-322, monitored by the Human Factors Operations Research Laboratories, Air Research and Development Command, Bolling Air Force Base, Washington, D. C. The writer is particularly indebted to Dr. George A. Miller for his advice and counsel during the course of the study and the writing of the report.

TABLE 1

WORDS USED IN THE EXPERIMENT

IA	IB	IIA	IIB	IIIA	IIIB
no	know	threw	through	corps	core
rain	reign	phrase	frays	one	won
paste	paced	seed	cede	scene	seen
whirled	world	ruff	rough	wail	whale
colonel	kernel	aloud	allowed	metal	mettle

nology, naive with respect to the purpose of the experiment. Each had been accustomed to tachistoscopic testing during a preliminary session. No word was shown in the preliminary session which was to be used in the main experiment.

Apparatus.—Word recognition thresholds were measured with an electronically timed mirror tachistoscope. This instrument was built by Dr. J. Volkmann of Mt. Holyoke College, who generously permitted its use in the M.I.T. laboratories. In design, it closely resembles the tachistoscope recently described by Kupperian and Golin (3). The timing circuit[1] allowed variation of the exposure duration in 17 steps from 60 to 290 msec. Under the illumination conditions employed, no *S* recognized any stimulus word at the shortest exposure time, while failures of recognition at the longest exposure occurred only three times.

The customary modified method of limits was used in the determination of thresholds. Measurements on each *S* were made from an individual base line—one of the 17 exposure speeds—which had been determined during the preliminary session. Each word was successively exposed at increasing exposure times above this base line until it was correctly recognized. Guessing was encouraged. Since the purpose of the experiment was the comparison of different classes of words for each *S* individually, accurate exposure times were not required. Only the ordinal number (1–17) of the threshold exposure was recorded, and only ordinal statistics are employed in evaluating the results.

Stimulus materials.—The words used as stimulus materials are shown in Table 1. They were arbitrarily divided into three groups of five homonym pairs each, as indicated.

Procedure.—Each *S* was instructed that he would be given 1 min. to study a card on which a list of ten words was typed. He was told that some of these words would appear among the words to be displayed in the tachistoscope, and that he should keep the list in mind as well as possible since this might help him to see or guess the words in tachistoscopic presentation.

The words shown to each *S* were chosen from the experimental materials in the manner indicated in Table 2. The order of the words was determined separately for each *S* from a table of random numbers.

After *S* had studied the card for 1 min., it was taken from him and the tachistoscopic test was begun. Thresholds were determined successively for 17 words, in the manner described above. The first two words were irrelevant practice items. The remaining 15 were drawn from the experimental materials according to Table 2. The order in which the words were tested was determined individually for each *S* from a table of random numbers.

The design of the experiment, as outlined in Table 2, provides certain controls. Each *S* was tested on three classes of words: (*a*) five words from the list he had been shown; (*b*) five words that were homonyms of the list words not directly tested; (*c*) five words that were new to him. These three classes will be referred to as set words, homonyms, and controls respectively. Every word in the experiment appeared equally often (twice) in each of the three classes. The words of Group IA, for example, were set words

TABLE 2

EXPERIMENTAL DESIGN

Subject Number	Words Presented on Preliminary List	Words Tested Tachistoscopically
1	IA, IIA	IA, IIB, IIIA
2	IB, IIB	IA, IIB, IIIA
3	IA, IIA	IB, IIA, IIIB
4	IB, IIB	IB, IIA, IIIB
5	IA, IIIA	IA, IIIB, IIA
6	IB, IIIB	IA, IIIB, IIA
7	IA, IIIA	IB, IIIA, IIB
8	IB, IIIB	IB, IIIA, IIB
9	IIA, IIIA	IIA, IIIB, IA
10	IIB, IIIB	IIA, IIIB, IA
11	IIA, IIIA	IIB, IIIA, IB
12	IIB, IIIB	IIB, IIIA, IB

[1] Designed by Dr. E. Davis.

for *Ss* No. 1 and 5, homonyms for *Ss* No. 2 and 6, and control words for *Ss* No. 9 and 10.

Results

It was not possible to make meaningful comparisons between the tachistoscopic test scores of the several *Ss*, because a different base line was employed for each *S*, as explained above.

In order to study differences among individual words, the data for each *S* were analyzed to find his median threshold. Each of the words on which he had been tested was then classified as falling either above, below, or at this median. This analysis revealed substantial differences in the perceptibility of the different words. Seven of the 30 words in Table 1 were never recognized more quickly than the median by any *S;* these were *phrase*, *scene*, *wail*, *metal*, *cede*, *core*, *mettle*. Five words were never slower than the median; these were *ruff*, *one*, *kernel*, *allowed*, *whale*. These differences are not related to the main variable of the experiment (set), since each word appeared equally often in each of the three experimental classes. To the extent that the differences among words are significant, they are probably due to differences in structural characteristics and in frequency of use.

Data from each *S* were analyzed to determine how many of each class of words were recognized faster than, slower than, or at his median recognition threshold. The results are totaled in Table 3. Table 3 shows that set words tend to be recognized more quickly than average, while homonyms and controls behave substantially alike in being slower than average. Although these results appear clear-cut, it is impossible to test their statistical reliability, since the 180 entries in Table 3 are not independent. There

TABLE 3

Comparison of Experimental Conditions

Recognition Speed	Set Words	Homonyms	Controls	Total
Slower than median	15	32	29	76
Faster than median	29	16	17	62
(At median)	(16)	(12)	(14)	(42)

Note.—The figures represent the total number of cases (all *Ss* and words) in the given categories.

are 6 different entries based on each word, and 15 based on each *S*.

The reliability of the results was therefore estimated by comparisons based on independent samples, in this case the 12 *Ss*. It was first necessary to determine for each *S* whether he saw homonyms more quickly than controls or the reverse, and to make similar determinations for the two other crucial pairings (homonyms vs. set words, set words vs. controls). Such determinations for a given *S* were made by finding the direction of the cross product in the appropriate fourfold table: a table with one entry for each homonym or control (say) recognized with a speed above or below the median threshold for the *S*.

A comparison based on these determinations revealed no significant differences between homonyms and controls. Five *Ss* were faster on homonyms and 4 on controls, while 3 were equally fast on each. Set words, however, were consistently superior to both other classes. Nine of 11 *Ss* recognized set words more readily than controls (one showing no difference), while 11 of 12 *Ss* recognized set words faster than homonyms. Simple sign tests indicate that the probabilities of obtaining results deviating from the chance 50% by at least the given amounts are .065 for 9/11 and .006 for 11/12. These probabilities are based on the null hypothesis that

every *S* is equally likely to recognize either class of words more readily; both tails of the binomial distribution are thus used.

Thus the superiority of set words to both homonyms and controls is reasonably reliable, while there is no discernible difference between homonyms and controls. This substantiates the hypothesis that the set does not merely raise the probabilities of certain verbal responses, but rather facilitates the perception of specific visual patterns.

Summary

Twelve *S*s were given a list of ten words to study for 1 min. Tachistoscopic duration thresholds were secured on five words from this list, five words which were homonyms of items on the list, and five control words. A balanced design was used. The results indicate that the preliminary presentation facilitated the recognition of specific items on the list, but in no way facilitated the recognition of their homonyms. Since the same verbal response is employed in reporting a homonym as in reporting the word itself, it appears that the effect of a set of this type is to facilitate recognition processes without generally facilitating the corresponding verbal responses. This conclusion may be related to certain recurrent problems in the interpretation of perceptual data.

References

1. Graham, C. H. Visual perception. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 868–920.
2. Howes, D. H., & Solomon, R. L. Visual duration threshold as a function of word-probability. *J. exp. Psychol.*, 1951, **41**, 401–410.
3. Kupperian, J. E., & Golin, E. An electronic tachistoscope. *Amer. J. Psychol.*, 1951, **64**, 274–276.
4. McGinnies, E. Emotionality and perceptual defense. *Psychol. Rev.*, 1949, **56**, 244–251.
5. Osgood, C. E. *Method and theory in experimental psychology*. New York: Oxford Univ. Press, 1953.

(Received August 20, 1953)

Journal of Experimental Psychology
Vol. 47, No. 6, 1954

EYE–HAND SPAN IN SIMPLE SERIAL TASKS [1]

E. C. POULTON

Applied Psychology Research Unit, Cambridge, England

It was aimed to discover the psychological factors determining the limits of the range of eye-hand span in simple serial tasks. The normal range was defined for a three-choice serial task, and attempts were made to obtain eye-hand span values outside this range, by modifying the experimental conditions in convenient related tasks. The aim contrasts with that of previous studies of eye-hand and eye-voice span, which have concentrated mainly upon the mean length of the span, and its variations with experience and experimental material (6, pp. 730–734).

METHOD

Three-Choice Double-Joystick Task

A 16-response three-choice task was selected for experimental study in order to reduce the chance of random responses being correct, while minimizing the effective number of choices at any one time. The *S* was presented with two circles side by side on a vertical screen, in each of which was an arrow pointing N, S, E, or W. He had two joystick keys in front of him, one for each hand, which he had to move in the directions of the arrows. The 16 different combinations of the 4 left-hand and 4 right-hand display directions were arranged in a sequence on a film strip to make a three-choice task of approximately uniform difficulty. Displays were alternately symmetrical, e.g., NN, WE, and asymmetrical, e.g., NW, ES; neither hand moved in the same direction twice running; and the same one of the 16 different combinations never occurred twice in three consecutive displays. The film was moved automatically one frame at a time, at speeds determined electronically. Responses were recorded graphically by four two-way electrically activated pens; an additional pen recorded the time at which the displays appeared.

Practice experiment.—This experiment was performed to determine the normal range of eye-hand span at fast and at slow speeds of response, and with relatively little and with greater amounts of practice, and to compare these values with the corresponding choice reaction times. Four *S*s received eight combined practice and experimental periods; another four *S*s received only six periods, owing to mechanical faults in the apparatus. They had two periods daily, one in the morning, the other in the afternoon, in each of which responses were made to 800 displays. The first 640 displays were always presented in series of 80, at about the fastest constant speed at which accurate performance was possible, with rest pauses between series. In all except the first and last periods the next 100 displays were presented as a single series, while the last 60 were split up into three short series of 20 displays each. These short series were presented at faster speeds, in order to determine the fastest speed at which short bursts of 20 errorless responses could be made.

In the first and last periods these three short series came immediately after the first 640 displays, and the last 100 displays were presented as five further short series, one under each of the following conditions: (*a*) 3.0 sec. per display, *S* to respond to each as quickly as possible (choice reactions in Table 1); (*b*) 1.25 sec. per display in the first practice, .75 sec. in the last (a slow speed for *S*); (*c*) .75 sec. per display in the first, .50 sec. in the last (a fast speed); and (*d*) two series with a display speed starting within *S*'s capabilities, but increasing until it was beyond them (rapidly accelerating displays in Table 4). This last condition, and the two next experiments, were designed to produce a transient increase in the length of the eye-hand span, for preliminary experiments had shown that the span tended to increase with display speeds which were slightly too fast, when performance was about to break down.

Response time and correctness were determined from the graphic records. The mean response times (Table 1) were calculated from the 10 middle displays of each series of 20, excluding the occasional errors. In the two rapidly accelerating series (Table 4), *E* only measured the response time of the next to the last correct response before errors and/or omis-

[1] The influence of Professor Sir Frederic Bartlett, F.R.S., and of Dr. N. H. Mackworth is clearly shown in this paper. The *S*s were supplied by the Royal Navy. Financial support from the British Medical Research Council is also gratefully acknowledged.

sions occurred, neglecting the occasional single error (mean display speed .47 sec. per display in the first practice, .33 sec. in the last). This response time was chosen because it was usually the longest which occurred while correct *serial* responses were being made. When making the last correct response before errors and/or omissions, there was no evidence from the graphic record that *S* was watching for the succeeding display. Thus it could not be concluded with certainty that he was still performing a serial task.

Too fast display speed experiment.—Before performing this experiment, 12 new *Ss* each responded to 500 displays, all but the last 80 in a standard practice period on the previous day. No exact equalization of the initial practices of the different experiments was possible, for except in the practice experiment a specific experimental procedure had to be learned during the practice, and in the remaining tasks (see below) the task also was different. In all except the practice experiment the practices were designed to be just long enough to teach *S* the required procedure. Practice and experimental periods did not exceed 1 hr. in length, additional periods being allocated to experiments taking longer than this.

In the experiment *S* had to respond to six series each of 80 displays, presented at speeds of 1.5, 1.0, .8, .63, .55, and .48 sec. per display. At one or more of the speeds in the middle of the range (mean of .60 sec. per display), *Ss* showed long series of correct responses followed by groups of errors and/or omissions. Two such series were selected from the graphic records of each *S*, and the response time of the next to the last correct response in each was measured.

Slow acceleration and extra display experiment.—Before this experiment another 12 *Ss* had each to respond to 600 displays, all but the last 120 in a standard practice period on the previous day. In one part of the experiment they had to respond to four series each of 80 displays. The display speed accelerated slowly over the first 70 items from 1.0 to .63 sec. per display, and rapidly over the last 10 items to .43 sec. per display. In each series the response time of the next to the last response of the last long group of correct responses was measured (mean display speed .63 sec. per display).

In a second part of the experiment each *S* had to respond to six series of displays, two series being presented at each of the three speeds of 1.2, 1.0, and .8 sec. per display. After every fourth normal display (except the first) an additional display was inserted midway between it and the normal fifth, in order to double the display speed in this region. Altogether, six insertions were made into each series, giving 12 displays which were followed by only half the normal time interval before the arrival of the next display. The response times to these displays were measured in all cases where both the response itself and the subsequent response were correct (17% of the responses themselves were wrong or omitted; 19% of the subsequent responses). Six *Ss* performed the slow acceleration half of the experiment first, the other six the extra display half.

Remaining Tasks

Three-choice joystick task.—This and the next experiment were designed in attempts to reduce the eye-hand span beyond its normal lower limit of length. An 8-response three-choice task was substituted for the 16-response task, so that each successive display item could be contained within a single circle. Twelve *Ss* were presented with a single horizontal line of equal-sized circles projected on a vertical screen, the circles being separated from each other by a distance equal to their radii. Filling each circle was an arrow which pointed in one of the eight main directions of the compass. Again the directions were arranged in a sequence on a film strip to give a three-choice task of approximately uniform difficulty. Display items were alternately "easy," e.g., N, E, and "difficult," e.g., NE, WS, and the same display direction never occurred twice in three consecutive display items. The whole display moved at a steady speed from right to left across the screen, circles appearing on the right and disappearing on the left. The *S* had to respond to the arrow in the circle in the extreme left-hand position on the screen, as each circle reached this position, by moving an 8-position joystick key in the corresponding direction. The exact position on the screen was marked by pointers above and below the line of circles. Responses were again recorded graphically by four two-way electrically activated pens. The *E* controlled a fifth pen by depressing a key, and in one version of the electric circuit *S*'s responses were only recorded if they occurred while this key was depressed.

Before the experiment *S* responded to 2,000 display items, all except the last 400 in three practice periods during the previous week. In the experiment the display moved at .6, .5, and .4 sec. per circle in different experimental periods. Four sizes of display were presented at each display speed, $\frac{3}{4}$, $1\frac{3}{4}$, $2\frac{3}{4}$, or $3\frac{3}{4}$ additional circles being visible ahead as *S* responded to the arrow in the left-hand circle. Seeing $\frac{3}{4}$ of a circle ahead was just sufficient to give the direction of the next arrow. The 12 combinations of three speeds and four sizes of display were arranged randomly in two alternative latin squares. One set was performed under free recording conditions, responses being recorded whenever they

occurred. The other set was performed under constrained recording conditions, responses only being recorded when *E* depressed his key, which he did every time a circle was in the extreme left-hand position on the screen. Under both conditions *S* heard a series of taps with which his responses had to synchronize, in addition to seeing the positions of the circles on the left of the screen. If he responded at the correct time under constrained recording conditions, he heard a bell instead of the tap. Conditions out of the two sets were performed alternately.

Responses had to be made to 53 display items under each condition, but only the last 50 of these were scored. Under constrained recording conditions it was possible to determine the percentage of correct responses correctly timed, and also the percentage correct but exactly one or more items too early or late (Table 2). Under unconstrained recording conditions it was possible to determine the over-all response deficit ("omitted" in Table 2) and the over-all percentage of wrong responses made. The remainder were taken to have been correct but to have been made slightly too early or late, so that they were not recorded under constrained recording conditions.

Harmonic tracking task.—Harmonic tracking was chosen because it is approximately a three-choice type of task (the three choices being to continue responding at the same speed, and to respond at increased or decreased speeds) in which perfect prediction is theoretically possible but is not achieved in practice (**3, 4**). A curved course was drawn on a paper tape, which moved at a speed of 1.0 in. per sec. Fifteen *S*s had to track this, using a ball-point pen which could be moved against a Perspex bar lying at right angles to the paper tape. In different experimental periods the course was open to view for 8.0, .4, .3, .2, and .1 sec. (in.) before it reached *S*'s pen. He always had an uninterrupted view of the paper tape carrying the course and his response line, once it had passed his pen.

Three course speeds were presented with each length of preview. The fast course was produced by two equal-sized harmonic displacements of 52.0 and 42.0 cycles per min., and a displacement twice this size of 21.0 cycles per min. The medium speed course was half the speed of the fast, the slow a quarter the speed. The total amplitude of movement was about 4.5 cm. The 15 combinations of three course speeds and five lengths of preview were arranged in a latin square, such that each *S* had all his experimental periods at one course speed before proceeding to the next. Three 1-min. practices were given just before the experiment: the fast course with an 8.0-sec. preview, the slow course with a .3-sec. preview, and the medium speed course with a .1-sec. preview, in that order.

In the experiment, tracking for 1.0 min. was performed under each condition. A sample of ten errors, measured in terms of time every 4.0 cm. (1.58 sec.) over the last quarter of each record, was taken. The mean lag was calculated by taking into account the signs of the errors (whether *S* was ahead of the course or behind it), but the signs were neglected in calculating the mean error (Table 3). In addition, a measure of jerkiness was obtained by counting the number of discrete visible changes in the speed of movement over the last half of each record that did not correspond to changes in the speed of movement of the course.

16-choice double-joystick task.—In this experiment and the next, instructions were given to increase the eye-hand span, display items having to be memorized and subsequently recalled. In this experiment the 16-response three-choice task was changed to a 16-choice task, in order to prevent the failure to recall one item from reducing the chances of the next one being recalled correctly. Eight *S*s were presented with the same display as in the three-choice double-joystick task, except that the 16 possible display combinations were completely randomized on a film strip. Responses were again recorded graphically. The *S* had to respond at lags ranging from 1 to 3. At Lag 1 he had to respond to the display which had just disappeared, at Lag 2 to the display before this, and at Lag 3 to the display one further back still. Five speeds of response were tried at each lag: 1.0, 2.0, 3.0, and 4.0 sec. per response, and response at *S*'s own speed. The mean speeds selected were 1.7, 2.0, and 2.2 sec. per response at Lags 1, 2, and 3, respectively.

Before the experiment responses were made to 850 displays at lags ranging from 0 to 3; all except the last 50 were presented in three standard practice periods on different days. In the experiment the 15 combinations of three sizes of lag and five speeds of response were given to each *S* in a different random order. Under each condition a series comprised 15 items, but only the middle 7 of them were scored.

Harmonic-step tracking task.—Harmonic-step tracking is a task of approximately three-choice type (the three choices being to move an average distance beyond the midline, or rather further or less far); yet it avoids the simple signal patterns which occur in a three-response three-choice task, because movements have to alternate on either side of the midline. A moving paper tape carrying a zigzag course could be seen through a 1.5-mm. slit lying at right angles to it. Twelve *S*s had to watch the course in the slit, and to note

the positions at which it stopped as it changed the direction of its movement. These positions had to be copied at lags ranging from 0 to 3, using a wire pointer attached to a ball-point pen. The pen moved in a parallel slit, which was screened by a mask also attached to the pen. The *E* watched *S* continuously, and told him when he was responding at a lag one greater or smaller than that required. Before the experiment 1,000 practice responses were made at lags ranging from 0–3, all except the last 300 in a practice period on the preceding day.

In different experimental periods three course speeds were used. The slow course was produced by one harmonic displacement of 13 cycles per min. and two displacements, each 2/5 as large, of 10.5 and 5.25 cycles per min. This gave alternate side-to-side movements at an approximately constant rate (2.3 sec. per movement), but of varying amplitude, the maximum being 4.5 cm. The medium speed course was twice as fast, and the fast course four times as fast as the slow course. The 12 combinations of three course speeds and four sizes of lag were arranged in a latin square, such that each *S* performed all lags at one speed before proceeding to the next. The experimental periods were equated for number of decisions required. With the slow course they lasted 60 sec. each, with the fast course 15 sec. The alignment error was measured in millimeters for ten consecutive responses, always selected from a similar part of the course.

Subjects

The 79 *S*s were all enlisted men in the British Royal Navy, aged between 17 and 31. None performed more than one experiment.

Results

Variations in the normal range of eye-hand span in the practice experiment with the three-choice double-joystick task are shown in Table 1. Only twice was the normal span as long as the time interval between displays; it was never as long as twice the time interval.

In this experiment the mean fastest display speed at which 20 correct responses could be made to a series of 20 displays was .70 sec. per display before practice, .50 sec. after.

The effects upon performance of reducing the preview from $1\frac{3}{4}$ to $\frac{3}{4}$

TABLE 1

Normal Range of Span in Three-Choice Double-Joystick Task

Seconds per Display	Response Times in Seconds			
	Before Practice		After Practice	
	Mean	Mean *SD*	Mean	Mean *SD*
3.00*	.44	.061	.34‡	.036‡
0.50			.32§¶	.037§
0.75	.48†	.069	.51	.045
1.25	.52	.066		

* Choice reactions.
† Different from both 1.25- and 3.00-sec. rate at the .05 level.
‡ Before practice—after practice $p = .01$.
§ Different from .75-sec. rate before practice at the .001 level.
¶ Different from .75-sec. rate at the .001 level and 3.00-sec. rate at the .05 level.

items in the three-choice joystick task are shown in Table 2. The speed of .5 sec. per display gave results lying between those for the speeds of .4 and .6 sec. The two longest previews ($2\frac{3}{4}$ and $3\frac{3}{4}$ items) gave no additional information apart from revealing a tendency to respond too early. With the .30-sec. preview about half the responses were correct, although the majority of these were late; the remainder were incorrect or omitted.

Table 3 shows the effects of reducing the size of the preview in harmonic tracking from 8.0 to .1 sec. The medium speed course gave results lying between those for the fast and slow courses. As the preview with the fast course was reduced from 8.0 sec., the first just significant[3] change in performance was an increase in lateness (lag), which was found with the .3-sec. preview. With .2 sec. the mean size of the error and jerkiness (as previously defined) had both increased sig-

[3] Significance levels beyond $p = .001$ will be called "highly significant," beyond $p = .01$ "significant," and beyond $p = .05$ "just significant." Both tails of the assumed normal distribution have always been used except where stated.

TABLE 2

EFFECT OF ATTEMPT TO REDUCE THE SPAN IN THREE-CHOICE JOYSTICK TASK

Measure	Preview with .40 Sec. per Display		Preview with .60 Sec. per Display	
	½-Item (.30 sec.)	1½-Item (.70 sec.)	½-Item (.45 sec.)	1½-Item (1.05 sec.)
Correct and correctly timed	18%*	43%§	60%	80%‡
Correct but 1 or 2 items too late	10%	4%§	6%	1%†
Correct but 1 or 2 items too early	2%*	2%	0%	3%
Correct but just out of time	20%	12%	15%	1%
Incorrect	34%	29%¶	19%	14%
Omitted	16%	10%¶	0%	1%

* Not significantly above a level of performance which could have been achieved by responding at random.
† ½-item—1½-item $p = .05$.
‡ ½-item—1½-item $p = .01$.
§ ½-item—1½-item $p = .001$.
¶ Incorrect and Omitted combined ½-item—1½-item $p = .001$.

nificantly, and all three measures showed a steep rise as the preview was further reduced to .1 sec.

When serial performance was about to break down because the display speed was slightly too fast, response times were found to increase. Three different ways of inducing this situation were tried, using the three-choice double-joystick task: (*a*) a display speed starting within *S*'s capabilities and increasing slowly or more quickly until it was beyond them; (*b*) a display speed slightly too fast for *S*, which could only be managed successfully during an initial spurt; and (*c*) a display speed within *S*'s capabilities, which was doubled intermittently by the insertion of an extra display. The percentages of correct responses made after the corresponding display had disappeared from view (i.e., at Lag 1) when performance was about to break down under these conditions are shown in Table 4, together with the number of observations made under each condition. The eye-hand span was never as long as twice the interval between displays during accurate serial performance.

Tables 5 and 6 show the effect of requiring *S* to respond with constant-sized lags in the 16-choice double-joystick and harmonic-step tracking tasks respectively. At Lag 1 he had to respond to the display which had just disappeared, at Lag 2 to the display before this, and at Lag 3 to the display one further back still. Omissions did not occur in step tracking, and made up only 1% of all responses in the 16-choice task.

TABLE 3

EFFECT OF ATTEMPT TO REDUCE THE SPAN IN HARMONIC TRACKING

Measure	Fast Course		Slow Course		Difference Required for $p = .05$
	.1-Sec. Preview	8.0-Sec. Preview	.1-Sec. Preview	8.0-Sec. Preview	
Mean lag (sec.)	.14	.02	.04	.01	.04
Mean error (sec.)	.16	.06	.10	.07	.03
Mean number of jerks in 30 sec.	46	30	37	32	6.0

TABLE 4

EFFECT OF ATTEMPTS TO INCREASE THE SPAN IN THREE-CHOICE DOUBLE-JOYSTICK TASK

Condition	Percentage of Correct Responses at Lag 1	*N*
Rapidly accelerating displays		
Before practice	40%	16
After practice	55%	16
Slowly accelerating displays	2%	48
Too fast constant speed displays	25%	24
Extra displays	9%	620

DISCUSSION

The lower limit of eye-hand span.—Both before and after practice in the three-choice double-joystick task, the shortest mean response time was obtained when *S* was responding about as quickly as he accurately could (Table 1). This time was in the region of the choice reaction time (Table 1; but see below).

Table 2 shows the effect of an attempt to shorten the response times in the three-choice joystick task. A time limit was inserted by which correct responses had to be made if they were to be counted as correct, and in certain periods the preview was reduced to ¾-item (.30 sec. at the speed of .4 sec. per display; .45 sec. at .6 sec. per display). This size of preview was chosen because it was just large enough to enable *S* to determine the direction of his next response, while making the previous one at the correct time. At .4 sec. per display under these conditions the percentage of correctly timed correct responses fell to the chance level, and there was an increase in late responses even at .6 sec. per display. There was thus no evidence that response times of .30 sec. or below could be obtained consistently in this three-choice task.

The following evidence suggests that the *amount of information required* (5) in order to respond accurately was a determinant of the lower limit, the minimum duration of the span varying directly with the information required. This conclusion is similar to one arrived at by Hick for choice reaction times (1).

1. The permitted orders of the 16 possible display items in the three-choice double-joystick task had to be learned before the task actually became three choice. Before learning was completed, *S* had to obtain, on the average, more information from each item before he could respond accurately, because his expectancy had not been reduced to the three alternatives actually possible. Hence his mean response time was longer before practice than after, when responding as quickly as he could after the presentation of each display (Table 1; 3.00 sec. per display, and .75 sec. per display before practice, .50 sec. after). Before practice 3.00 sec. per display was more favorable for short response times than faster display speeds, presumably because the additional time before the presenta-

TABLE 5

EFFECT OF INCREASED SPAN IN 16-CHOICE DOUBLE-JOYSTICK TASK

Time Available per Response	Percentage of Correct Responses		
	Lag 1	Lag 2	Lag 3
2.0 sec. or more	98%*	[43%]†	[25%]
1.0 sec.	70%	[27%]†	[9%]

Note—The bracketed values are not significantly above a level of performance which could have been achieved without memory.
* 2.0 sec. different from 1.0 sec. at .001 level.
† Lag 1—Lag 2 $p = .001$.

TABLE 6

EFFECT OF INCREASED SPAN IN HARMONIC-STEP TRACKING TASK

Time per Response	Mean Error (mm.)			
	Lag 0	Lag 1	Lag 2	Lag 3
2.3 sec.	1.1	1.8	3.5*†	[4.4]
1.2 sec.	1.6	2.3	[5.0]†	[6.5]
0.6 sec.	3.8	[4.5]	[6.1]†	[7.1]

Note.—The bracketed values are not significantly above a level of performance which could have been achieved by a constant-sized movement.
* 2.3 sec. different from 1.2 sec. at .01 level and from .6 sec. at .001 level.
† Lag 0—Lag 2 $p = .001$.

tion of each display enabled *S* to eliminate certain alternatives. But after practice the additional time was a slight disadvantage.

2. In tracking the slow harmonic course, which was in theory completely predictable, little information had to be obtained from the display before responding, once the course was familiar. Hence reducing the preview from 8.0 to .1 sec. resulted in a mean lag of only .04 sec., an increase which was just significant only when using a one-tailed *t* test (Table 3). Adding this mean lag to the duration of the preview (.1 sec.) gives a mean response time of only .14 sec., which is less than the mean simple reaction time of about .17 sec. for a visual signal (6, p. 324). Short response times in harmonic tracking without any preview, and the nature of the advance information which can be used, have been studied previously (3, 4).

3. The speed of the fast harmonic course was four times that of the slow, so that in tracking it more information had to be obtained from it per unit time. Hence the mean lag with the .1-sec. preview was longer (Table 3). The increase in mean error was due to the increased lag. The increased jerkiness (increase in the number of unnecessary changes in the speed of movement) was due to an increase in incorrect or only partially correct anticipations, which subsequently required correction.

The upper limit of eye-hand span.— Both before practice and after, the mean response time in the three-choice double-joystick task was longer at slow display speeds than at fast (Table 1). However, the normal span was seldom as long as the time interval between displays, except when performance was about to break down because the display speed was slightly too fast (Table 4).

At least two distinct factors determined the upper limit. First to be considered is *the amount of information to be held in store* between perception and response. The *S* could respond without storing information, provided his delay was less than the time interval between displays (Lag 0 in Table 6). When responding at Lag 1, the information corresponding to one response had to be stored, at Lag 2 that corresponding to two responses, and at Lag 3 that corresponding to three. Tables 5 and 6 show that with instructions to increase the lag, and therefore the amount of information to be held in store, there was a corresponding decrease in the accuracy of performance. This explains why when *S* had to respond accurately in the three-choice double-joystick task at speeds within his capabilities, he rarely let his responses fall behind as far as Lag 1, never as far as Lag 2.

A second factor is *the rate of storing and reproducing information.* Tables 5 and 6 show that with the lag, and therefore the amount of information to be held in store, kept constant, increase in display speed reduced the accuracy of performance. Speeding up the display reduced the time available for storing and reproduction, and also reduced the time for which the information had to be held in store. The former must have been responsible for the result found, for variations in storage time of the order used in these experiments (up to 6.0 sec.) do not appear to affect the accuracy of recall (2).

A combination of these two factors explains the findings in the three-choice double-joystick task when performance was about to break down because the display speed was slightly too fast. The *S* never responded at Lag 2 during accurate serial performance, and often did not even respond at Lag 1 (Table 4).

Summary

The normal range of eye-hand span was defined for a three-choice serial task, and attempts were made to increase and reduce the span in related tasks, in order to discover the factors determining its limits.

The normal span was shortest at fast display speeds. As there were 16 alternative responses in the three-choice task, the length of the span under these conditions depended upon learning which sequences were possible and which were not. An attempt to reduce the span in an 8-response three-choice task to .30 sec. was unsuc-

cessful; it simply produced an increase in errors and omissions. However, a comparable method gave a mean response time of only .14 sec. in tracking a slow harmonic course. The mean response time was longer and performance more jerky when the course was faster.

The normal span was longest at slow display speeds, but was seldom as long as the time interval between display items. It could be increased temporarily by speed stress, although never as far as twice the interval between displays during accurate serial performance. When the span was increased beyond this point by instructions, errors increased, especially with faster speeds of performance.

It was concluded that the amount of information required in order to respond was a determinant of the lower limit of the span. The amount of information to be held in store, and the rate of storing and reproducing information, were both determinants of the upper limit.

References

1. Hick, W. E. On the rate of gain of information. *Quart. J. exp. Psychol.*, 1952, **4**, 11–26.
2. McLane, A. S., & Hoag, J. E. The curve of forgetting in the first three minutes. *Amer. J. Psychol.*, 1943, **56**, 105–110.
3. Poulton, E. C. Perceptual anticipation in tracking with two-pointer and one-pointer displays. *Brit. J. Psychol.*, 1952, **43**, 222–229.
4. Poulton, E. C. The basis of perceptual anticipation in tracking. *Brit. J. Psychol.*, 1952, **43**, 295–302.
5. Shannon, C. E., & Weaver, W. *The mathematical theory of communication.* Urbana: Univ. of Illinois Press, 1949.
6. Woodworth, R. S. *Experimental psychology.* New York: Holt, 1938.

(Received August 21, 1953)

Journal of Experimental Psychology
Vol. 47, No. 6, 1954

TRANSFER OF TRAINING AS A FUNCTION OF DEGREE OF RESPONSE OVERLEARNING [1]

GEORGE MANDLER

Harvard University

The majority of studies investigating transfer of training as a function of degree of prior learning have paid relatively little attention to the problem of whether or not *S*s are trained on previously familiar or unfamiliar responses, e.g., whether or not the assembly of discrete response units into more complex responses takes place during the early stages of training. Thus, Siipola and Israel (6) pretrained their *S*s on motor responses before introducing the actual training situation. Underwood (8) used adjectives, i.e., highly overlearned and familiar verbal responses. Bruce (1), on the other hand, using nonsense syllables, restricted the original learning situation to relatively low degrees of overlearning. The present study will investigate the effect of degrees of training on three transfer situations patterned after Bruce's paradigm (1). The responses employed consist of novel assemblies of discrete response units. Degree of *over*learning will be stressed; i.e., transfer effects will be related to practice on the training task beyond the stable establishment of stimulus-response connections or habits. A detailed theoretical analysis of transfer effects associated with (*a*) *S*s' learning how to perform the response per se—the integration of the response—and (*b*) high degrees of overlearning has been presented elsewhere (4).

METHOD

Task and apparatus.—The task consisted of operating hand switches in sequences of three at a time on a switchboard having six switches arranged in a hexagon; the stimuli were letters of the alphabet. No time limit was set for the performance of the correct response after the presentation of the stimulus, and *S* could engage in free trial and error behavior in discovering each correct response. All *S*s, except one control group, were trained on an original task (Task I) and tested on a transfer task (Task II).

Figure 1 shows the apparatus. The main units were: (*a*) A *response board* (15 in. square) with six Mallory spring-return lever switches, 6 in. apart, arranged in a regular hexagon. (*b*) A *stimulus screen* (6 × 12 in.) on which the stimuli, eight letters of the alphabet, were produced by arrangements of lights behind translucent glass. The letters were 7½ × 4 in. at approximately 35 in. from *S*. When no stimuli were presented, the screen was an undifferentiated white surface. (*c*) The *success signals* were presented in the three 2-in. circular openings below the stimulus screen (numbered 1, 2, and 3 in Fig. 1).

The required sequences of letter stimuli, correct levers (in the proper order), and success signals were automatically programmed in fixed series of 40 stimulus presentations (ten trials of original task and five trials of transfer task) on a 4-pole, 22-position stepping relay. The left success signal was lighted when the first correct lever was pressed, and the center and right signals were lighted when *S* pressed the second and third correct levers, respectively. Incorrect responses did not alter the problem setting or the signals. No activation of the second, or third, success signal was possible unless the preceding lever, or levers, had been pressed. The interval between the lighting of the third success signal and the presentation of the next letter on the stimulus screen was .9 sec.

A complete record of all responses made and their latencies was obtained on an Esterline-Angus recorder, Model AW.

Subjects.—The *S*s were 60 volunteers from introductory courses in psychology, mostly college sophomores. They were assigned at random

[1] This paper is part of a dissertation submitted to the Graduate School, Yale University, in partial fulfillment of the requirements for the Ph.D. degree. The writer wishes to thank Drs. F. D. Sheffield, C. I. Hovland, G. A. Kimble, and R. P. Abelson for their help at various stages of the study. Special thanks are due Dr. L. H. Beck for his valuable help in the design and construction of the apparatus.

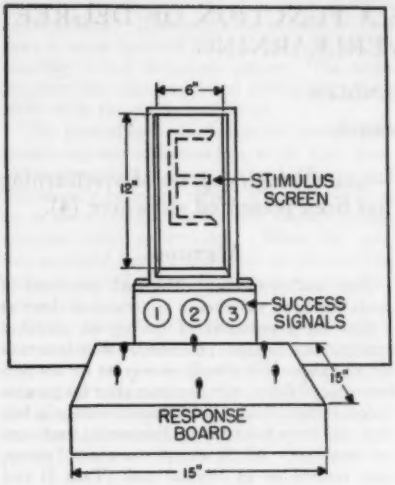


FIG. 1. The apparatus

to ten groups (five degrees of training and two forms) as they appeared for the experiment.

Instructions.—The *S* was shown the eight stimuli, asked to name the letters, and then told the response requirements and the requirements that had to be satisfied to light the success signals. He was urged to solve each problem as quickly as possible, but if he asked if speed was involved in the task, he was informed that it was not a speed test, but that he should try to keep working consistently at the problem without letting pauses occur. Finally, *S* was informed that in the course of the experiment *E* might stop him to make some adjustments in the apparatus.

Transfer conditions.—The *S*s learned a four-item training task (Task I) to the various degrees of overlearning and then were switched to an eight-item transfer task (Task II). Each "item" consisted of a letter stimulus and a three-lever response sequence. Task II simultaneously tested transfer effects for four conditions of transfer. The stimuli and responses used in Task I will be called "old" stimuli and responses; those first introduced in Task II will be called "new." This results in four transfer conditions:

I: Learning to make a new response to an old stimulus.

II: Learning to make an old response to a new stimulus.

III: Learning to make an old response to an old stimulus when these have not been previously paired. This involves the recombination of stimuli and responses used in Task I.

IV: Learning to make a new response to a new stimulus. This condition was used as a control for such phenomena as the warm-up effect demonstrated by Thune (7).

Two forms (A and B) of Tasks I and II were used. Table 1 shows the stimuli and responses used in both forms of the two tasks.

The response sequences fulfill the following conditions: (*a*) Any subsequence of two levers that was used in a given response was not used in any other response. (*b*) In Task I each lever was used twice; in Task II each lever was used four times. (*c*) In those transfer conditions (I and III) where two interfering responses are elicited by the same stimulus, the two response sequences have only one common lever, and that lever occurs in a different position in the sequence of the Task II response than it does in the Task I response.

A "trial" in Task I consisted of a repetition of all four stimulus-response pairs. Ten different orderings of pairs within trials were used, the only restriction on the ordering being that the first pair of a succeeding trial was different from the last pair of the preceding trial. This gave a total sequence of 40 pairs (10 trials) which was successively repeated. A comparable method was used with the transfer task except that with eight pairs per trial, five trials completed a run-through.

Degree of training conditions.—Preliminary experiments had shown that at levels of training of less than ten errorless trials the transfer phenomena were extremely variable; therefore, the smallest amount of training given was ten errorless trials. The degrees of training used in the present experiment consisted of 0, 10, 30, 50, and 100 errorless trials on Task I. A different group of *S*s was used for each of these degrees of training; these five groups will hereafter be designated in terms of these amounts of training.

Group 0 had no training on Task I. In Group 10, *S* was stopped on Task I as soon as his performance included 10 errorless trials. In Groups 30 and 50, *S* was stopped for 1.5 min. after achieving 10 errorless trials, and then was given additional blocks of trials including at least, and not more than, 10 errorless trials, to a total of 30 and 50 errorless trials respectively. These blocks were separated by 1.5-min. intervals. In Group 100, *S* was treated the same as were Group 50 *S*s except that when he had finished the last block of trials, giving him 50 errorless trials, he was told to come back the next day at the same time. On the second day, he was given an additional five blocks of 10 errorless trials each on Task I.

TABLE 1

LETTER STIMULI AND LEVER SEQUENCE RESPONSES USED IN THE TWO FORMS OF TASKS I AND II

Transfer Conditions Represented in Task II	Form A				Form B			
	Task I		Task II		Task I		Task II	
	S	R	S	R	S	R	S	R
I. (Old S–New R)	O H	3–6–5 6–2–1	O H	2–4–6 4–1–5	C U	5–4–2 1–3–4	C U	3–2–6 5–3–1
III. (Old S–Old R)	F L	1–3–4 5–4–2	F L	3–6–5 1–3–4	E A	3–6–5 6–2–1	E A	1–3–4 3–6–5
II. (New S–Old R)			C U	5–4–2 6–2–1			F L	6–2–1 5–4–2
IV. (New S–New R)			E A	5–3–1 3–2–6			O H	4–1–5 2–4–6

Note.—For the purpose of this presentation, levers are numbered 1 through 6 in a clockwise order starting with the lever in the 12 o'clock position.

After the conclusion of the requisite number of errorless trials, *S* was given a 3-min. rest while *E* changed the problem on the wiring board to Task II. The *S* then worked on Task II to a criterion of two successive errorless trials (16 correct responses) or 20 trials, whichever occurred last. Group 0 proceeded, after identical initial instructions, directly to Task II.

RESULTS

Training results.—An analysis of variance was performed on the training data (i.e., on Groups 10, 30, 50, and 100) using trials to the first perfect trial and frequency of correct responses in achieving ten perfect trials. As expected, there were no significant differences in the randomly selected groups, but there was a significant difference between Forms A and B by the former measure. However, no significant differences were found for the two forms beyond 30 correct trials. Table 2 shows the means and *SD*'s for both forms.

If mean time to achieve successive blocks of trials containing 10 correct trials is used as the measure, there was no improvement in performance beyond 50 correct trials; if mean errors are used, there was no improvement beyond 40 correct trials. This is shown in Fig. 2, in which both forms are combined and in which all four training groups are combined for the first point; Groups 30, 50, and 100 are combined for the second and third point; and so on. It will be seen that the time measure increases following the 24-hr. interval (in Group 100), followed by a recovery to the level at the end of the first day. However, errors did not increase over the 24-hr. interval. Since there was no observable improvement in Task I performance after 50 correct trials, overlearning, defined above as training beyond

TABLE 2

MEANS AND *SD*'S OF NUMBER OF TRIALS AND FREQUENCY OF CORRECT RESPONSES UP TO THE FIRST CORRECT TRIAL ON TASK I

Form	Trials to First Correct Trial		Correct Responses to First Correct Trial	
	Mean	*SD*	Mean	*SD*
A	31.4	13.6	32.6	20.3
B	23.3	8.9	23.5	11.3

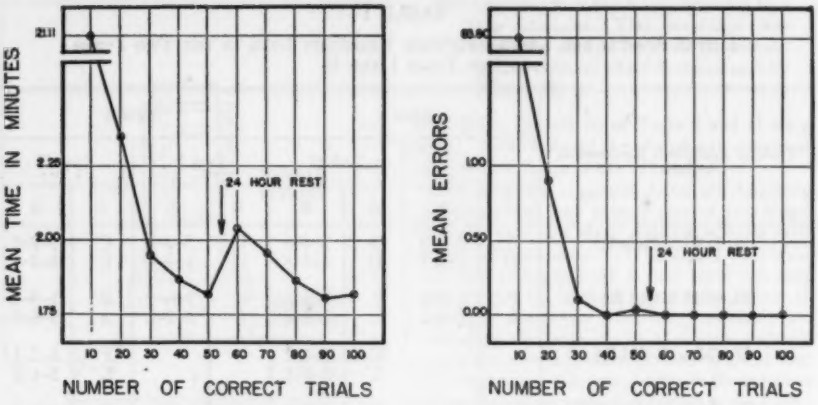


FIG. 2. Mean time and errors to perform successive blocks of ten perfect trials, shown as a function of level of training (number of correct trials achieved)

the establishment of stable S–R connections, was presumed to have been accomplished.

Transfer results.—Two different transfer measures were obtained for each *S* on the three transfer conditions; Cond. IV was used as the control. One measure, referred to hereafter as the F-20 score, was the raw transfer minus control difference in the frequency of correct responses during the first 20 transfer trials on Task II. The other measure, referred to as the FCR score, was the logarithm of the ratio of transfer to control performance in terms of the number of the trial on which the first correct response occurred. The logarithmic transformation in the latter measure provides for symmetry around zero transfer scores. These measures, using *S*s as their own controls, also utilize control conditions (Cond. IV) that are specific to the particular form used, i.e., all *S*s in the Form A group were given transfer scores based on Cond. IV in Form A, and similarly for the Form B group.

Analyses of variance of the transfer measures showed no significant difference between the two forms; the data were therefore combined. The analyses did show significant effects due to transfer conditions ($p < .01$), as well as significant interactions between degree of training and transfer conditions ($p < .05$). The specific p values presented below refer to the significance of t tests based on mean scores.

Figures 3 and 4 show the results for the two transfer measures, F-20 and FCR. As expected, Group 0, without previous training, showed some differences between the four conditions (in this case differences in difficulty of the S–R pairs). In order to take account of these initial differences between the S–R pairs used, the data were corrected to assume a common origin (zero transfer). Comparisons within any one transfer condition are unaffected by this correction. Statistical comparisons between transfer conditions among training groups were controlled for initial difficulty by testing the significance of differences in all such comparisons against the differences obtained in the corresponding comparisons with Group 0.

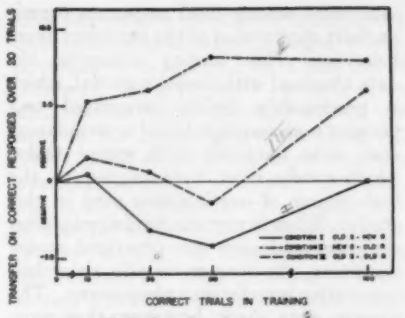


FIG. 3. Transfer of training as a function of degree of original training in terms of frequency of correct responses in 20 transfer trials (F-20)

In learning a new response to an old stimulus (Cond. I), the F-20 and the FCR measures show increasing negative transfer up to Groups 30 and 50, respectively. After these points, less negative transfer was found. Only the FCR measure, however (Fig. 4), approaches significance in these trends, the p value for 0 vs. 30 being .07, and for 30 vs. 100 being .06.

In learning an old response to a new stimulus (Cond. II), both measures show increasing positive transfer as degree of original learning increases. The reversal on the FCR measure for Group 30 (Fig. 4) is not significant ($p = .26$). Examination of the data shows that this reversal is primarily due to two *Ss* who showed extraordinarily fast acquisition in the control and slow acquisition in the transfer condition. Comparison of Groups 0 and 100 is significant ($p = .01$ on FCR, $p = .02$ on F-20). Actually, some *Ss* in Group 100 exhibited one-trial learning under this condition; i.e., the correct response was made and maintained from the second trial on.

Learning to make an old response to an old stimulus when these have not been previously paired (Cond. III) shows no significant transfer effects until the highest degree of overlearning on the original task is reached. Not only Group 0, but also Group 50, is significantly different from Group 100 ($p < .05$ for both measures).

At the highest degree of training, Cond. III shows significantly more positive transfer than Cond. I in Group 100 ($p < .05$ for both measures). This difference is apparent in Fig. 3 and 4. It can also be noted that Cond. II is superior to both Cond. I and III. The superiority of II over III is significant ($p < .05$) both for 10 and 50 correct training trials. The superiority of Cond. II over I is significant ($p < .05$) at all points.

In terms of the pattern as a whole, it is clear that the order from positive transfer toward negative transfer is Cond. II, III, and I, regardless of degree of training or the significance tests at individual points. In Fig. 3 or 4, the 12 points (four degrees of training and three conditions of transfer) could have, by chance, ordered themselves in $3!^4 = 1296$ different ways. Only one of these is the perfect result in Fig. 3, which therefore has a probability value of $\frac{1}{1296}$ or .0007. One reversal, as in Fig. 4, can occur in eight ways (two for any of the four groups), making the probability $\frac{8}{1296}$ or .006 of getting this particular pattern.

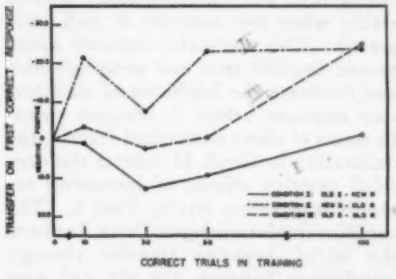


FIG. 4. Transfer of training as a function of degree of original training in terms of attainment of the first correct response (FCR)

Discussion

The increasing facilitation on Cond. II (New S–Old R) as a function of degrees of training bears out Wylie's generalization (9) that transfer effects are positive when an old response can be transferred to a new stimulus. This effect is also shown in Bruce's data (1) when they are corrected for the positive transfer found in his New S–New R condition—the warm-up control in the present experiment.

The lack of negative transfer in Cond. I (Old S–New R) when the degree of overlearning was high is similar to the effect found by Underwood (8) with familiar verbal material. His data showed positive transfer with high degrees of learning, but included no controls for warm-up effects.

The absence of observable transfer in Cond. III (Old S–Old R) under low degrees of overlearning is followed by a rise in the positive direction with high overlearning. This substantiates Siipola and Israel's (6) finding with a similar task—codes on a telegraphic key.

These findings bear out a theoretical analysis of transfer (4) that postulated three major variables: (*a*) The integration of the response, which refers to *S*'s learning how to perform the response. Parts of a complex response serve as stimuli for each other, and repeated reinforced repetitions increase the integration. (*b*) Stimulus-response association as a function of reinforced repetitions of the response in the presence of the stimulus. (*c*) A symbolic counterpart or analogue of the response, which is most stable when the response is well integrated. This postulated variable would permit implicit trial and error behavior and facilitate the inhibition of an incorrect response before it becomes overt. In terms of these theoretical factors, the facilitation in Cond. II reflects the one-sided positive effects of increasing response integration during Task I. The curvilinear relationship in Cond. I reflects the initial negative transfer through interference between the old and new responses, but the ultimate reduction back to zero interference when highly integrated wrong (old) responses permit implicit elimination of the incorrect overt behavior. This finding connects the data obtained with verbal material, which is presumably highly integrated and permits such implicit denial or avoidance, with data obtained with motor tasks, which rarely have been carried to the high degree of overlearning used in this study. When integrated verbal responses are compared with less practiced motor responses, the former usually show less associative interference phenomena. The present data show, however, that even with motor responses negative transfer due to associative interference can be eliminated, if the response is highly overlearned during training. The findings for Cond. III reflect, for lesser degrees of training, the summation of positive gains from the presence of some response integration and of negative effects of response interference. With high degrees of overlearning, however, the interference tends to be eliminated and response integration becomes the dominant factor.

Gagné, Baker, and Foster (2) have recently made a prediction directly applicable to Cond. III (Old S–Old R). This condition exemplifies what they have called "stimulus-response reversal," for which they predict maximum amounts of negative transfer. Porter and Duncan (5) have substantiated this prediction in a *verbal* paired-associate task. The present data, however, show less negative transfer in this condition than in Cond. I (Old S–New R). While it may be argued that in Cond. III the response used in Task II has some degree of similarity with the response used in Task I (one identical lever used in a different position of the sequence), and thus may counteract negative transfer effects, the same similarity holds for Cond. I. Thus, it appears that Gagné, Baker, and Foster's prediction is not generally applicable.

A final word should be said about the relation between the theoretical approach from the response-learning standpoint and the stimulus approach of Gibson (3). Her model of verbal learning, based on differentiation and discrimination on the stimulus side, is not designed to deal with

response factors. While she does deduce increasing generalization followed by differentiation as learning on a first list is increased, she specifically exempts those situations in which identical stimuli are used in the second list (here Cond. I). There could be little argument that the overlearning effects arose from *S*s' learning to differentiate the stimuli used (letters of the alphabet), which were deliberately selected to minimize stimulus differentiation. Also, stimulus differentiation was not even possible as a factor in Cond. II since new stimuli were used in the transfer task.

Summary

Sixty *S*s learned a motor task (pressing sequences of levers arranged in a spatial pattern) in response to letters as stimuli. The degrees of prior training were 0, 10, 30, 50, and 100 errorless trials on the training task, with 12 *S*s in each of these groups. The transfer task simultaneously tested four different conditions, three transfer and a fourth condition which controlled for warm-up. The findings were:

1. Learning to make a new response to an old stimulus showed an initial increase in negative transfer followed by a return to zero transfer for the highest degree of overlearning of the training task.

2. Learning to make an old response to a new stimulus showed increasing positive transfer as degree of original training was increased.

3. Learning to make an old response to an old stimulus, when these had not been previously paired, showed no significant effects at low degrees of training, but significant positive transfer at the highest degree of overlearning.

The relationship of these data to a theoretical framework (4) dealing with response factors in human learning and overlearning was discussed.

References

1. Bruce, R. W. Conditions of transfer of training. *J. exp. Psychol.*, 1933, **16**, 343–361.
2. Gagné, R. M., Baker, K. E., & Foster, H. On the relation between similarity and transfer of training in the learning of discriminative motor tasks. *Psychol. Rev.*, 1950, **57**, 67–79.
3. Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.
4. Mandler, G. Response factors in human learning. *Psychol. Rev.*, in press.
5. Porter, L. W., & Duncan, C. P. Negative transfer in verbal learning. *J. exp. Psychol.*, 1953, **46**, 61–64.
6. Siipola, E. M., & Israel, H. E. Habit interference as dependent upon stage of training. *Amer. J. Psychol.*, 1933, **45**, 205–227.
7. Thune, L. E. The effect of different types of preliminary activities on subsequent learning of paired-associate material. *J. exp. Psychol.*, 1950, **40**, 423–438.
8. Underwood, B. J. Proactive inhibition as a function of time and degree of prior learning. *J. exp. Psychol.*, 1949, **39**, 24–34.
9. Wylie, H. H. Transfer of response in the white rat. *Behav. Monogr.*, 1909, **3**, No. 5.

(Received August 22, 1953)

Journal of Experimental Psychology
Vol. 47, No. 6, 1954

RESISTANCE TO EXTINCTION AS A FUNCTION OF VARIATIONS IN STIMULI ASSOCIATED WITH SHOCK [1]

HOWARD MOLTZ

New York University [2]

The anxiety reduction hypothesis (**2, 8, 9, 10, 11**) states that cues accompanying noxious stimuli come to evoke a drive state, which when so elicited is termed "anxiety." Any response successful in removing these cues will reduce or eliminate anxiety, with the result that the tendency to make that response will be strengthened in accordance with the principle of reinforcement. After shock is no longer administered, the response tendency will be reinforced as a consequence of anxiety reduction following escape from the cues to which anxiety had been conditioned.

A number of studies (**3, 4, 7**) have revealed that responses occurring in close temporal contiguity with shock acquire unusual strength so that they persist even when no longer leading to reward. Such persistent nonadaptive behavior has been termed "fixation." Farber (**4**) has maintained that it is possible to explain the phenomenon by the learning principles involved in the anxiety reduction hypothesis. After no shock is administered, the persistence of the response might be due to secondary reinforcement following the reduction of anxiety. That is, if anxiety had been conditioned to cues associated with shock, escape from these cues would result in secondary reinforcement of the escape response. Farber reasoned that if the persistence of the response was a consequence of uncontrolled secondary reinforcement, feeding in a situation in which anxiety had developed should reduce or eliminate the anxiety and thus reduce or eliminate the source of reinforcement. Thus, if feeding at the locus of shock was shown to decrease resistance to extinction, the hypothesis that response fixation was a consequence of uncontrolled secondary reinforcement would be supported. Farber's results indicated that feeding at the locus of previous shock application was "therapeutic" in that those *Ss* not fed at the locus of shock required significantly more trials to reach the extinction criterion of two successive reversals of the original maze habit than those *Ss* who were fed.

Following from the anxiety reduction hypothesis and the Farber study is the implication that if anxiety is conditioned to cues accompanying the application of shock, then therapy (feeding or placing at the locus of previous shock application) designed to reduce anxiety will do so to the degree to which it contains or preserves those cues. That is, the extent to which therapy reduces anxiety will be related to the similarity in stimulus conditions between shock and therapy; the greater the similarity, the fewer the number of trials to reach a criterion of extinction. If it is shown that a difference in stimulus conditions between shock and therapy increases the num-

[1] This article is based on a portion of a dissertation submitted to the department of psychology of New York University in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The writer wishes to express his sincere appreciation to Professor Frank N. Marzocco for generous advice and criticism throughout all phases of this investigation.

[2] Now at the University of Illinois.

ber of responses to extinction, then the hypothesis that therapy is more effective in reducing anxiety if it contains the cues to which anxiety had been conditioned would be supported.

Method

Subjects and apparatus.—The *S*s were 96 experimentally naive, male albino rats of the Wistar strain. The ages of *S*s ranged from 50 to 75 days at the beginning of the experiment.

The apparatus was two single-unit T mazes, identical except that one was painted white and the other black. The roof of each maze was covered with .25-in. mesh hardware cloth. Cloth curtains, the same color as the maze, were placed in front of the doors to the goal boxes. Guillotine doors of light sheet metal were operated by *E*. The floor and sides of the left goal box of each maze were covered with .25-in. mesh hardware cloth; the right goal boxes remained unlined. The use of these different cues was intended to accelerate learning by decreasing stimulus generalization from one side of the maze to the other.

There was a distance of 40 in. from the outer end of the left goal box to the outer end of the right goal box. The length of the stem was 14 in. Each maze contained a starting box with a movable end wall designed to enable *E* to push the *S*s forward into the stem. Brass rods of .25-in. diameter covered the floor of the choice point, arms, and stem of each maze. In the arms and choice point, the distance between the rods was .25 in., while in the stem it was .5 in. The grids in the arms and choice point could be activated independently of those in the stem. The 110-v., 60-cycle alternating current was adjusted by an autotransformer. The circuit included a 340,000-ohm resistor placed so that it would be in series with *S*.

Two straight alleys were also used; one black and the other white. Each straight alley was 38 in. long and 4 in. wide. The roof of each was covered with .25-in. mesh hardware cloth.

Preliminary training.—Prior to training, all *S*s were given five days of habituation to the feeding regimen of the experiment. During these five days, *S*s were fed an unlimited amount of wet mash once daily for 1 hr.

The *S*s were assigned at random to the black or to the white maze. The preliminary training consisted of 15 min. of maze exploration in groups of two or three *S*s, followed by four trials in a straight alley to food. The color of the straight alley in which any *S* was run was the same as the color of the maze in which it was subsequently run throughout the experiment.

Learning series.—Beginning on the second day, all *S*s were given 10 trials a day (except on the last day of the learning series) for the remainder of the experiment. At the beginning of each day's run, *S*s were under a 20-21-hr. hunger drive. The intertrial interval was always at least 10 min., with a modal interval of 17 min.

The side of the maze in which any *S* was to receive food (consisting of pellets weighing approximately .150 gm.) was determined randomly. For the remainder of the learning series food appeared only on that side. Runs to the incorrect side were not reinforced either by the presence of food or by the presence of the food container. Responses to the correct or incorrect side did not involve any delay. Correction or retracing after an incorrect response was not allowed. Throughout the learning series, the *S*s were run in the maze in which they had received their preliminary training. All *S*s were given a total of 105 trials.

After 40 trials *S*s were assigned at random to two groups. The *S*s of the control group were not shocked; those of the experimental group were shocked. A total of 48 *S*s was assigned to each group; 24 *S*s were run in the black maze and 24 in the white maze.

Shock could not be avoided merely by a reversal of choice since the grids on both sides of the choice point were activated. In a preliminary study it was found that the onset of strong shock was in some cases sufficient to effect a disruption of the previously learned response, so that these *S*s continued to run incorrectly throughout the shock trials. To prevent the disruption of the response learned during the first 40 trials, the door leading to the incorrect alley was closed for the experimental *S*s during the first five shock trials (Trials 41–45). The door leading to the incorrect alley was also closed for the control *S*s during Trials 41 to 45.

Shock was progressively increased during the course of the shock trials from an initial intensity of 125 v. to a final intensity of 136 v. Since the experimental *S*s soon refused to approach the shock grid, it was necessary to administer an occasional shock in the stem of the maze. The time allowed for a choice before shock was given in the stem was reduced from 5 to 3 sec. during the course of the shock trials and served to equate roughly the running times of the experimental and control *S*s.

During the course of the shock trials, the end wall of the starting box was moved forward to force the *S*s out. Both the experimental and control *S*s were allowed to remain in the starting box for 5 sec. before the end wall was moved forward. Moving the end wall forward prevented *S* from avoiding shock either by refusing

to leave the starting box or by running back into it.

Therapy procedure.—All *S*s were given therapy in the maze immediately after the last trial of the learning series. The area of the maze in which they were given therapy consisted of the choice point and the arm leading to the correct goal box. The *S*s were put in the area for two 5-min. periods, separated by approximately 1 hr. During the two 5-min. periods, one-half of the control *S*s and one-half of the experimental *S*s were fed. The remaining *S*s of both groups were placed in the area but were not fed. For the *S*s fed during therapy, pellets totaling 2 gm. were placed in the arm leading to the correct goal box. The placed and fed groups were divided into two additional subgroups. The *S*s of one subgroup were given therapy in a maze whose color was different from that in which the learning trials had taken place. The *S*s of the other subgroup were given therapy in a maze whose color was the same as that in which the learning trials had taken place.

The *S*s were also assigned to two other subgroups. For *S*s of one subgroup the grid was absent during therapy; for *S*s of the other subgroup the grid was present. A thin strip of beaver board was placed over the floor of the area for those *S*s for whom the grid was absent.

Extinction trials.—About 24 hr. after therapy, the extinction trials were begun. For each *S* the food reward was shifted to the previously incorrect side of the maze, and no shock was administered. During the extinction trials all *S*s were run in the same maze in which their preliminary training and learning trials had taken place.

All *S*s were run for ten trials a day, with 10 min. or more between trials. The extinction procedure was continued until *S* reached a criterion of two successive responses to the now correct (but previously incorrect) food side. As in the learning series, retracing after an incorrect response was not permitted.

Results

Learning series.—In the initial 40 trials of the learning series the experimental and control *S*s showed no significant difference in the mean number of correct responses. An examination of the number of correct responses during the last 65 trials of the learning series reveals a mean difference significant at the .05 level of confidence. The mean number of correct responses for the experimental and control *S*s was 64.14 and 62.0, respectively. However, if two extreme scores in the control group are disregarded, this mean difference becomes insignificant. For no reason apparent to *E*, these two *S*s did not consume the food reward during the last 20 to 25 trials of the learning series. During these 20 to 25 trials they responded at a chance level. If only the experimental *S*s are considered, no significant difference was found between the various subgroups. In all cases, between-groups variance was smaller than within-cells variance.

Extinction series.—In order to test the experimental hypothesis, it was first necessary to demonstrate that the conditions employed were adequate to produce fixation. An analysis of variance using original extinction scores did not seem desirable since the requirement of homogeneity of variance could not be met. However, transforming the original extinction scores by taking the common logarithm of each score satisfied the requirement. Table 1 shows the mean of each subgroup before carrying out the transformation.

TABLE 1

Mean Number of Trials to Extinction for Each Group

Comparisons	Experimental *S*s ($N = 48$)	Control *S*s ($N = 48$)
Placed	33.70	5.58
Fed	18.00	6.29
Black	21.38	5.54
White	30.33	6.33
Same	15.92	6.58
Different	35.79	5.29
Grid	22.42	6.29
No grid	29.29	5.58
Grand mean	25.85	5.94

TABLE 2

SUMMARY OF ANALYSIS OF VARIANCE FOR THE NUMBER OF TRIALS TO EXTINCTION FOR ALL *Ss*

Source	*df*	Mean Square	*F*
Experimental, control	1	5.35	35.42**
Placed, fed	1	.044	<1.00
Black, white	1	.007	
Same and different maze colors	1	.354	2.21
Grid, no grid	1	.013	
Experimental, control × same and different maze colors	1	.853	5.66*
Within cells	64	.151	

* Significant at the .05 level.
** Significant at the .01 level.

Table 2 shows the analysis of variance for the experimental and control *Ss* combined. Table 3 reveals the analysis of variance for the experimental group *Ss* separately. Only the main effects and the significant interactions are included in each table.[3]

Table 1 shows that the experimental *Ss* required a greater mean number of trials to reach the extinction criterion than the control *Ss*. The *F* shown in Table 2 indicates that the difference is significant at the .01 level of confidence. There can be little doubt that the conditions employed in the present experiment were adequate to produce fixation.

An examination of the extinction scores of all *Ss* combined revealed no significant difference between *Ss* run in the white maze and *Ss* run in the black maze. The extinction scores of the experimental *Ss* separately and those of the control *Ss* separately also showed no significant difference. All interactions involving black and white proved insignificant. The results indicate that the color of the maze in which the preliminary training and learning trials took place did not significantly affect rate of extinction.

Table 1 shows that for experimental *Ss* that had different maze colors in the shock and therapy situations, the mean number of trials required to reach the extinction criterion was 35.79; for experimental *Ss* that had the same maze colors, the mean number of trials to reach the extinction criterion was 15.92. The *F* shown in Table 3 indicates that the difference is significant at the .05 level of confidence.

Before the hypothesis is confirmed with respect to the variable being considered, it is necessary to examine the extinction scores of the control *Ss*. Table 1 shows that for the control *Ss* that had different maze colors in the shock and therapy situations, the mean number of trials required to reach the extinction criterion was 5.29; for control *Ss* that had the same maze colors, the mean number of trials required to reach the extinction criterion was 6.58. This difference is

TABLE 3

SUMMARY OF ANALYSIS OF VARIANCE FOR THE NUMBER OF TRIALS TO EXTINCTION OF THE EXPERIMENTAL GROUP *Ss* SEPARATELY†

Source	*df*	Mean Square	*F*
Placed, fed	1	.255	1.20
Black, white	1	.001	<1.00
Same and different maze colors	1	1.14	5.36*
Grid, no grid	1	.083	
Within cells	32	.213	

* Significant at the .05 level.
† Interactions were not significant.

[3] The complete analysis of variance for the number of trials to extinction of the experimental group *Ss*, the control group *Ss*, and the experimental and control group *Ss* combined has been deposited with the American Documentation Institute. Order Document No. 4217 from the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $1.25 for 35 mm. microfilm or $1.25 for 6 by 8 in. photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

not significant. Since, in addition, the interaction between the presence and absence of shock on the one hand, and same and different maze colors on the other, was significant, it is concluded that a change in maze color did not significantly increase the number of trials to extinction of the control *Ss*. The results support the hypothesis that a change in stimulus conditions between shock and therapy did increase the number of extinction trials of the experimental *Ss*.

For one-half of all *Ss*, the grid was absent during therapy; for the remaining half, the grid was present. The experimental *Ss* for whom the grid was absent did not require significantly more trials to extinguish the previously learned habit than did the experimental *Ss* for whom the grid was present. When the extinction scores of both the experimental and control *Ss* are combined, the interaction between shock–nonshock and grid–no grid also proves insignificant.

One-half of the experimental *Ss* and one-half of the control *Ss* were fed during the two 5-min. therapy sessions; the remaining *Ss* were simply placed in the situation but were not fed. Analysis of the data revealed that feeding and placing did not have significantly different effects on the rate of extinction. All interactions involving feeding and placing were likewise insignificant.

Discussion

Since the results indicated that fixation did occur, it was possible to test the experimental hypothesis that the effectiveness of therapy in reducing anxiety would vary as a function of the degree to which it contained or preserved the cues present at the time of shock. An analysis of the extinction scores of experimental *Ss* that had different maze colors in the shock and therapy situations, as compared with *Ss* that had the same maze colors, supported the hypothesis.

Not only do the results support the experimental hypothesis, but also the anxiety reduction hypothesis; for if it is assumed that anxiety is a learned response-produced drive conditioned to cues accompanying shock, then any condition reducing anxiety should be most effective in the presence of the cues evoking it. Consequently, if *Ss* were again subjected to a situation in which they had been shocked previously, but without shock being administered, these cues would come to evoke less anxiety. As a result, the maintenance of the response by means of secondary reinforcement would be decreased. However, if the cues during therapy differed from the cues to which anxiety had been conditioned, less anxiety reduction could be expected to occur. Since, for the *Ss* of the latter group, the stimulus conditions under which anxiety was reduced were not identical with those evoking it, the *Ss* should be more anxious when placed in the original situation.

It is a well-established phenomenon that, just as positive habits manifest stimulus generalization, extinction effects will also manifest this tendency (6). The data obtained by Hovland (5), and Bass and Hull (1), show that extinction will decrease as the extent of the difference between the evoking stimulus and the stimulus originally conditioned increases. It is reasonable to expect that if anxiety were conditioned to the cues of a white maze, therapy in a black maze would be less effective since reduction of anxiety would not be direct, but would be a result of the generalization of extinction effects from black to white.

While the hypothesis was supported in terms of the difference in the number of trials required for extinction between experimental *Ss* whose maze colors during shock and therapy were different, as compared with experimental *Ss* whose maze colors were the same, the hypothesis received no support when the grid and no-grid groups were examined.

The following interpretation is offered for the lack of significance found between the grid and no-grid groups. The *S*s of the present experiment lived and were fed in a cage having a thick wire mesh floor that permitted the dropping of fecal matter. Since the grids of the maze were spaced quite closely, it is possible that the tactual stimulation experienced in the living cage was similar to that of the maze. As a result, eating and drinking responses associated with the cage floor might have been elicited by the maze grids. If these responses interfered with the anxiety responses also conditioned to the maze grids, then either no anxiety, or less intense anxiety, would be evoked by the grids as compared with the maze color. If the tactual stimulation received from the grids was not an anxiety-eliciting cue, or was one of minimal value, its presence or absence would not be expected to affect significantly the subsequent rate of extinction.

No significant difference was found in the number of trials to extinction between *S*s who were fed during therapy and *S*s who were simply placed but not fed. Placing is assumed to have been effective in reducing anxiety, because anxiety responses appeared and were not followed by shock—the situation being one that forced the responses to occur since *S* could not escape. Failure to reward the anxiety response resulted in its being extinguished. In addition to failing to reward, feeding might have had the effect of forcing *S* to perform responses incompatible with anxiety. Analysis of the data revealed that while *S*s fed in the therapy situation required somewhat fewer trials to reach the criterion of extinction than those simply placed but not fed, the difference was not significant at the .05 level of confidence. The lack of significance indicates that either feeding had no additional effect and that reduction or elimination of anxiety during therapy was due to simple experimental extinction, or that feeding did have the additional effect of forcing *S* to perform responses incompatible with anxiety, but that a larger number of *S*s and more extreme conditions would have been needed for it to have resulted in a significant decrease in the number of trials to extinction. The latter alternative is, of course, testable by using a larger number of *S*s, by using more extreme conditions, and by testing for spontaneous recovery of anxiety with varying times between therapy and retest in the maze.

Summary

Following from the anxiety reduction hypothesis and the Farber study (4) is the implication that if anxiety is conditioned to cues temporally contiguous with the application of shock, then therapy (feeding or placing at the locus of previous shock application) will reduce anxiety to the degree to which it contains or preserves those cues.

To test the hypothesis, 96 experimentally naive, male albino rats were used. All *S*s were given 105 trials in a single-unit T maze, with food reward in one of the goal boxes. During the last 65 trials, the experimental *S*s were shocked at the choice point and arm of the maze. The *S*s of the control group were not shocked. Immediately following the last trial, one-half of the *S*s were fed at the locus of shock; the remaining one-half were placed at the locus of shock but were not fed. The placed and fed groups were divided into two additional subgroups in which therapy was given in a maze whose color was, or was not, different from that during the learning trials. For one-half of all *S*s the grid was absent; for the other half the grid was present during therapy. Approximately 24 hr. after the therapy situation, the food reward was shifted to the previously incorrect side of the maze. All *S*s were run without shock until they reached the extinction criterion of two successive responses to the now correct (but previously incorrect) food side.

The conditions employed in the present experiment were adequate to produce fixation in the experimental (shock) groups. No significant difference was found between the extinction scores of experimental *S*s whose therapy situation contained a grid and those whose therapy situation did not; but when the maze color in the therapy situation was different from the maze color in the shock situation, the number of trials to extinction increased significantly. Since the extinction scores of the comparable control *S*s did not differ significantly as a function of a change in maze color, the experimental hypothesis was supported.

References

1. Bass, M. J., & Hull, C. L. The irradiation of a tactile conditioned reflex in man. *J. comp. Psychol.*, 1934, **17**, 47–65.
2. Dollard, J., & Miller, N. E. *Personality and psychotherapy: an analysis in terms of learning, thinking and culture.* New York: McGraw-Hill, 1950.
3. Everall, E. E. Perserveration in the rat. *J. comp. Psychol.*, 1935, **19**, 343–369.
4. Farber, I. E. Response fixation under anxiety and non-anxiety conditions. *J. exp. Psychol.*, 1948, **38**, 111–131.
5. Hovland, C. I. The generalization of conditioned responses: III. Extinction, spontaneous recovery, and disinhibition of conditioned and of generalized responses. *J. exp. Psychol.*, 1937, **21**, 47–62.
6. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
7. Kleemeier, R. W. Fixation and regression in the rat. *Psychol. Monogr.*, 1942, **54**, No. 4 (Whole No. 246).
8. Miller, N. E. Studies of fear as an acquirable drive: I. Fear as motivation and fear-reduction as reinforcement in the learning of new responses. *J. exp. Psychol.*, 1948, **38**, 89–101.
9. Miller, N. E. Learnable drives and rewards. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 435–472.
10. Miller, N. E., & Dollard, J. *Social learning and imitation.* New Haven: Yale Univ. Press, 1941.
11. Mowrer, O. H. A stimulus-response analysis of anxiety and its role as a reinforcing agent. *Psychol. Rev.*, 1939, **46**, 553–565.

(Received August 24, 1953)

Journal of Experimental Psychology
Vol. 47, No. 6, 1954

TEMPORAL NUMEROSITY: III. AUDITORY PERCEPTION OF NUMBER

PAUL G. CHEATHAM AND CARROLL T. WHITE

U. S. Navy Electronics Laboratory

The first paper of this series reported a study in which visual perceived number was investigated as a function of the objective number of flashes presented and the flash rate (1). The perceived number was found to be dependent on both the objective number of flashes and the total time in which they were presented. Stimulus flashes tended to be grouped subjectively into perceptual units which were consistent for all *Os*. The subjective number of flashes was determined primarily by the total time of a train of flashes rather than by the number of flashes. The upper limit for perceived number was about one subjective flash per 100 msec. of time, and the more stimulus flashes presented during this 100 msec. (the higher the flash rate), the closer this limit was approached.

The limiting of perceived number observed in these studies could have been produced by some temporal process characteristic of retinal action, or by some higher, more central, process. The problem of locus of such a temporal process was approached by recording human electroretinograms (ERG) during stimulation of the retina by flashes of light such as were used in the first study (6). The ERG records showed that the electrical activity of the light-adapted retina responded to every flash presented, up to rates of at least as high as 45 flashes/sec., and that the number of electrical responses of the retina did not bear any direct relationship to the perceived number of flashes. One may conclude from these results that the limitation of perceived number is not due to the action of the retina alone, but is produced primarily by some higher process.

Having ruled out the retina as a limiting mechanism for visual perceived number, and having hypothesized a higher central process as the determining factor, one might now look for the effects of such a central process in other sense modalities. The auditory modality seems especially suited for this second approach to the question of locus, and it has, in fact, been studied by other investigators, but from a different point of view (3, 5). If a similar temporal process is found to limit auditory numerosity, then the case for a central locus of this process will have been strengthened. The purpose of this study, then, was to repeat the visual research using sound pulses, in an attempt to discover whether a temporal limiting process can be found in a sense modality other than vision.

METHOD

The method was similar to that used for the study of visual perceptual number (1), except that sound pulses were presented instead of light flashes. This was accomplished by the use of a photronic tube and its associated circuitry. The tube was placed in the light path at the position of the original test field so that it now received the flashes of light. Photocurrents were amplified and used to key a 1000-cps tone from a beat frequency oscillator. The light beam activating the phototube was interrupted by sector disks, producing square pulses of 11-msec. duration at rates of 10/sec., 15/sec., and 30/sec. Rates were changed by changing the silent interval between pulses. The loudness of the steady, uninterrupted tone was held constant at 70 db.

The *O* listened to the pulses through earphones provided with doughnut cushions. A ready signal was given 2 sec. before each series of pulses was presented. The *O* was instructed to report the number of pulses heard. The method of constant stimuli was used for presenting the sound pulses; thus, for any given rate, the number of pulses was randomized, and the rates, in turn, were randomized. The data were taken during two experimental sessions, separated by approximately five months. At the first session each number of pulses was presented ten times at all rates to four *O*s who had observed in the visual studies. The second session was an exact replication of the first using the same *O*s.

Results

The reports of each *O* for the two experimental sessions were highly reliable. For this reason, all data have been grouped together and entered in Table 1. The perceptual numerosity functions are shown in Fig. 1. These curves show that: (*a*) the *O*s underestimated the stimulus number in all of the cases for which errors were recorded; (*b*) for each rate, estimation of stimulus number appears to decrease in a linear fashion; and (*c*) the variability of responses (Table 1) increases suddenly for mean perceived numbers higher than five.

Perceived number has been plotted as a function of time in Fig. 2. The abscissa shows cumulative time from the beginning of a train of pulses, with each point indicating a pulse. This set of curves shows that the auditory perception of number is definitely a function of the total elapsed time. In both Fig. 1 and 2, straight lines have been fitted to those data for which the perceived number does not agree with the number presented. The slopes of these straight lines correspond to the rates of 9/sec. for the 10/sec. and 15/sec. objective rates, and 11/sec. for the 30/sec. objective rate. Apparently, then, the subjective rate for

TABLE 1

Distribution of Responses to Varying Numbers of Pulses at Various Rates of Presentation

Number of Pulses	30/sec. Rate										15/sec. Rate								10/sec. Rate						
	Perceived Number										Perceived Number								Perceived Number						
	1	2	3	4	5	6	7	8	9	10	1&2	3	4	5	6	7	8	9	1&2	3	4	5	6	7	8
1	80																								
2		80									80								80						
3		59	21									80								80					
4		27	44	9								29	51							10	70				
5		6	55	19								10	49	21							17	63			
6			52	20	7	1							20	58	2						5	18	57		
7			13	54	9	4							9	34	33	4						3	26	51	
8			7	50	19	4								24	35	19	2						8	22	50
9			3	30	40	5	2							11	37	18	14								
10			1	8	59	10	2								20	30	25	5							
11				4	53	16	4	3							3	27	38	12							
12				2	50	20	7	1																	
13					35	26	6	7	6																
14					18	35	14	3	9	1															
15					12	29	21	7	8	3															
16						35	20	10	7	8															
17						13	33	17	5	12															

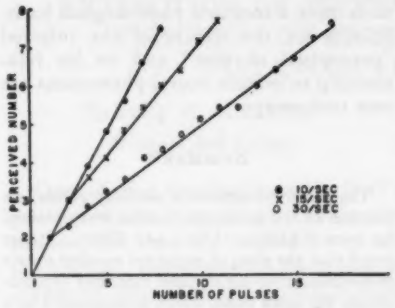


FIG. 1. Number of pulses perceived as a function of the number of pulses presented

auditory perception of sound pulses approaches a limit of about 9 to 11 pulses/sec.

DISCUSSION

The present study was undertaken to determine whether the auditory perception of discrete stimuli would show a temporal limiting effect such as that found in visual perception. When the present data for the three rates of presentation are plotted against time (Fig. 2), the resulting curves present a convincing argument for the presence of such a limiting process, more so than the corresponding visual curves. It is seen that consistent slopes are maintained, indicating a range of perceptual rates of only 9–11 pulses/sec., corresponding to a range of objective rates from 10–30/sec. With the visual curves, the limiting slope corresponded to a rate of about ten perceptual units per second, but this slope was not maintained over the whole range of responses. The difference in absolute value of perceptual number for the curves in Fig. 2 is likely to be misleading unless it is realized that this is brought about chiefly because, with the faster rates, there are more pulses available in a given time at the beginning of a series, when accurate counting is possible. Of primary importance, of course, is the similarity of the slopes of the three curves.

The task of reporting the number of sound pulses heard was found to be much more difficult than the reporting of perceived flashes in the visual counterpart of this study. This subjective appraisal was borne out by the results, which show a great deal more variability than that found with vision. The auditory data can be represented best by linear functions, while the visual data could not. There were so many points in the visual data which had no variability that a fitted curve would not have given a true picture of the results.

Marked differences in results should be expected between two such basically different sense modalities as vision and audition. Two major differences are apparent in our results: (*a*) the fusion exhibited during the first 100 msec. by the vision curve, and (*b*) the long plateau in the vision curve which begins around 300 msec. On the basis of additional work on this topic (6), it can be stated with a high degree of certainty that the fusion effect is due to the action of the rods during this period. This initial delay in the rise of the vision curve can be practically eliminated either by adapting the eye to a much higher intensity level before testing or by using a red stimulus light, or both.

The long plateau in the vision curve still awaits explanation. There is reason to believe that this is not a result of a purely visual process. An examination of the 30/sec. auditory data in Table 1, for example, shows a definite tendency for the modes to level off in the same

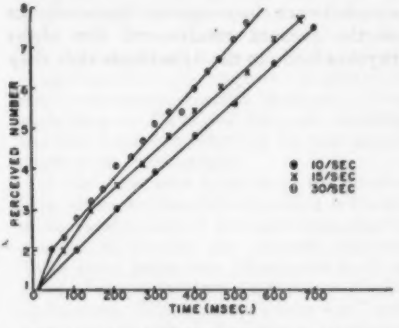


FIG. 2. Number of pulses perceived as a function of time

region of 9–13 pulses. Because of the skewed distribution of responses, this tendency is not so evident when the mean values are plotted. Further work is necessary before anything definite can be said about the basis of the phenomenon. If these differences are considered, the similarity of the upper limits of subjective rate for vision and audition seems even more significant.

The question of the peripheral or central locus of the limiting process for audition must now be considered. All available evidence would indicate that the peripheral mechanism is not responsible for this limiting process (2, 4). The action potentials in the auditory nerve will reproduce the stimulus tone up to 2,000–3,000 cps. The higher neural centers are capable of maintaining synchronized responses to stimulus tones up to frequencies of 1,000 cps, but at the highest center, the auditory cortex, the frequency of a pure tone is no longer reproduced. The cortical potentials, however, will follow a series of sharp clicks up to a frequency of about 100/sec. On the strength of this evidence, one can safely assume that the limiting does not take place at the receptor level.

The marked similarity of the upper subjective rates for both vision and audition strengthens the case for a common temporal process that limits and orders the perceptual events of both modalities. One can only speculate at present as to the nature of the mechanics of such a process; however, the close correspondence between the temporal characteristics of the present results and the alpha rhythm leads to the hypothesis that they both have a common physiological basis. Studies on the nature of the inferred "perceptual rhythm" and on its relationship to certain neural phenomena are now underway.

Summary

The perceived number of auditory pulses as a function of the presented number was obtained for rates of 10/sec., 15/sec., and 30/sec. It was found that the plots of perceived number versus presentation time for all three rates had approximately the same slope, which corresponded to a perceptual rate of about 10/sec. Since this was in basic agreement with visual results, it was concluded that a common temporal process in the central nervous system acts upon both modalities.

References

1. Cheatham, P. G., & White, C. T. Temporal numerosity: I. Perceived number as a function of flash number and rate. *J. exp. Psychol.*, 1952, **44**, 447–451.
2. Davis, H. Psychophysiology of hearing and deafness. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 1116–1142.
3. Garner, W. R. The accuracy of counting repeated short tones. *J. exp. Psychol.*, 1951, **41**, 310–316.
4. Geldard, F. A. *The human senses*. New York: Wiley, 1953.
5. Taubman, R. E. Studies in judged number: I. The judgment of auditory number. *J. gen. Psychol.*, 1950, **43**, 167–194.
6. White, C. T., Cheatham, P. G., & Armington, J. C. Temporal numerosity: II. Evidence for central factors influencing perceived number. *J. exp. Psychol.*, 1953, **46**, 283–287.

(Received August 31, 1953)

Journal of Experimental Psychology
Vol. 47, No. 6, 1954

EXTINCTION OF INSTRUMENTAL BEHAVIOR AS A FUNCTION OF FRUSTRATION AT VARIOUS DISTANCES FROM THE GOAL [1]

HARVEY M. ADELMAN AND GERALD ROSENBAUM
Michigan State College *Wayne University*

One of the implications of Hull's goal gradient hypothesis (2) is the proposition that the resistance to extinction of instrumental behavior is inversely related to the distance from the goal at which frustration occurs. This deduction follows from the assumption that instrumental action sequences frustrated close to the goal accrue more secondary reinforcement than those frustrated further away.

Lambert and Solomon (3) recently obtained a confirmation of this hypothesis by demonstrating that the running responses of rats in a straight runway extinguished more rapidly when blocking occurred further from the goal. Additional confirmation with human *Ss* is desirable, because the point at which frustration is introduced may prove to be an important variable in the social learning conditions producing individual differences in withdrawal.

The present study was designed to test the hypothesis by blocking the instrumental acts of human *Ss* at three different distances from the goal. The general procedure was to train three groups of *Ss* to perform an ordered series of five response patterns on a three-choice reaction time apparatus. Group A was then frustrated on Pattern 2, Group B on Pattern 3, and Group C on Pattern 4. Differences between the three groups in number of attempts to reach the goal, speed of response on the first pattern, and increases in speed following frustration served as measures of differential goal-directed tendencies.

[1] This article is based on the master's thesis of the first author (1), conducted under the direction of the second author at Wayne University. The second author is responsible for the manuscript in its present form.

METHOD

Subjects.—Thirty-nine introductory psychology students from Wayne University, 24 males and 15 females, were assigned to three equal experimental groups. As far as *E* could determine, all *Ss* were naive concerning the purpose of the experiment.

Apparatus.—The major unit of the apparatus consisted of a 20 × 28-in. baseboard into which were inserted three standard telephone switchboard plugs equidistant from one another and the sides of the board. These three plugs formed apertures into which a standard telephone jack, ordinarily resting in a starting aperture located 8 in. in front of the center plug, could be placed. A 7-w. stimulus lamp was located 4 in. behind the center plug and served as the reaction signal.

The stimulus-presentation unit was constructed so that *E* could illuminate the stimulus lamp by depressing any one of three switches connected to the three apertures on the baseboard. The *S* could extinguish the stimulus lamp by removing the jack from the starting position and plugging it only into the aperture corresponding to the switch *E* had thrown. The latency of *S*'s responses was recorded to the nearest .01 sec. by a Standard Electric Timer. The timer and stimulus lamp were activated simultaneously by any one of *E*'s switches, which were located behind a screening panel. This circuit could be broken only if the jack was plugged into the correct aperture. Return of the jack to the starting position illuminated a 7-w. ready lamp on *E*'s side of the panel, notifying him that *S* was in position for the next presentation of the stimulus lamp.

A row of five neon signal lamps was installed 28 in. above the base of the apparatus on *S*'s side of the panel to keep *S* informed concerning his progress in learning the response sequence. These signal lights were illuminated by *E* by means of a separate series of switches on *E*'s side of the panel. White numerals from 1 to 5 were painted below each of the lights in appropriate order.

A C. H. Stoelting Company stylus maze was used to permit *Ss* to succeed in a task that followed the main portion of the experiment. Total time on the maze was recorded by a stopwatch reading in .1 sec.

Prefrustration training trials. All *Ss* were seated and instructed as follows:

"This is a test of perceptual motor coordination similar to that employed by the Air Force. You will operate it in this manner. When this light, which we call the stimulus light, is illuminated, you will take the jack from the start position and plug it into one of these three positions. If the light still remains on, you will remove it from the position and plug it into one of the remaining two choices, and so on until the light goes off. When the light goes off, it means you have found the correct position, so remember your choice and return the jack to the start position and wait for the next trial to begin. Remember, in each case you will be required to locate the correct position since the stimulus light will remain on until you have done so.

"The lights will be presented in a series of four trials, each of which is in a definite pattern which you are to learn as fast as possible. Every time you get the pattern correct, that is, when you have turned off the stimulus light four times, the signal light above here will be illuminated. Now then, there are five signal lights, which means you must learn five patterns, in each of which you will have to turn out the stimulus light four times. Thus, to achieve a perfect score you will have to turn off the stimulus light with your first placement of the jack a total of 20 successive times. Remember, in order to learn where to place the jack you will have to concentrate on the five patterns corresponding to the five signal lights on the board. One way which you may find useful in remembering the correct patterns is to think of the openings in terms of left, center, or right. In this way you can later recall a particular pattern in terms of the positions of the openings. For example: one series of four trials may be center-center-left-left [demonstrate]. Any questions? Let's begin."

The five patterns used in this experiment were as follows: First, LCRL; second, RRRR; third, CCCC; fourth, LLLL; and fifth, CCCC. The clock used to time individual reactions was not reset after each trial, but the time was allowed to accumulate until one pattern (four trials) was finished. Number of errors per pattern was also recorded. Thus, for the complete problem there were five recorded time scores and five error scores, one for each signal light on the panel. When *S* had achieved three errorless performances on all five patterns, the criterion was reached and the test trials were begun.

Frustration trials.—The *Ss* were matched on the basis of both sex and number of trials to reach the criterion and assigned to one of the three groups. Thus the groups were each composed of eight males and five females, and were approximately equal in regard to number of training trials to reach the criterion. The following instructions were then given to all groups:

"Now that you have learned the correct patterns, I am going to give you a test. I am going to time each pattern and run you against the clock. Your speed will be compared with that obtained by college students throughout the country. You will be expected to be at least as rapid as the average college student. If, at the end of any pattern, your cumulative time is slower than that of the average college student, I will tell you to start at the beginning again. If you find that you are unable to achieve the required speed, please inform me at the end of any trial."

Although each of the three experimental groups received identical treatment in the training stage, frustration was introduced at a different point for each group during the testing trials. Group A was told that it was too slow after only one signal light had been illuminated (i.e., at the end of Pattern 2), Group B after two lights, and Group C after three lights. Following each frustration, *Ss* were asked if they would like to try again. The *Ss* in each group were frustrated at exactly the same point on succeeding trials. If *S* made any allusion to the number of trials he should attempt, he was told that he should continue until he either reached the goal of five lights in the average time, or until he thought he could not reach the mark and wanted to quit.

In addition to the time records for the learning stage, a time score was recorded for each pattern allowed on the frustration trials. Error scores and number of trials until *S* stated that he no longer wished to continue were also recorded. Following the testing period, *S* was told to try the maze, which was another measure of perceptual motor coordination and timed in the same manner as the apparatus. These time scores were recorded in .1 sec., and all *Ss* were informed that they had done very well. Systematic verbal reports were then taken concerning *Ss'* satisfaction, confidence, fatigue, boredom, and interest during the experiment by means of a five-point rating scale.

Results

The results of the experiment are summarized in Table 1 and the principal findings are presented in Fig. 1.

TABLE 1

STATISTICAL ANALYSES OF DATA FROM THE THREE EXPERIMENTAL GROUPS

Response Measures	Groups						*F*
	A		B		C		
	Mean	*SD*	Mean	*SD*	Mean	*SD*	
Trials to extinction	6.85	3.25	10.69	8.37	13.54	5.16	
Trials to extinction (transformed scores)	2.55	.59	3.07	1.13	3.58	.85	4.04*
Latency on Segment 1	3.76	.41	3.59	.49	3.38	.64	1.59
Improvement on Segment 1†	.66	.62	.90	.64	1.04	.58	1.21
Maze latency	13.11	3.32	13.93	3.69	12.42	2.38	0.58

* Significant beyond the .05 level of confidence.
† This measure refers to the difference in scores between the median of the last three training latencies and the median of the test trials on Segment 1 for each individual.

Since, with the exception of the segment at which frustration was introduced, the test conditions were constant for all *S*s, it was possible to combine the results from the three experimental groups into a three-point curve showing the variation of response strength as a function of frustration at different distances from the goal. The mean number of trials attempted by Groups A, B, and C were used directly in determining the ordinate values, while the values along the abscissa correspond to the segments at which Groups A, B, and C were frustrated.

Although all three groups took between five and six trials to reach the criterion of learning, it is apparent from both Fig. 1 and the top row of Table 1 that resistance to extinction increases progressively as the point at which frustration is introduced approaches the goal. In view of the heterogeneous variances of the trials-to-extinction measure, these scores were transformed to their square roots, and this trend was tested by means of simple analysis of variance. The second row of Table 1 reveals that the trend was significant at better than the .05 level of confidence. Further analysis of this trend by means of *t* tests employing the error term from the analysis of variance showed that the extreme points on the curve (Group A and C means) were significantly different at better than the .01 level, but that the means of Groups A and C did not differ significantly from that of Group B. The analysis thus indicates a monotonic trend in the predicted direction.

Further inspection of Table 1 reveals that strength of response, as measured by both the median response latency and the difference scores between the median latency on criterion trials and latencies on the first segment, corresponded with the trend obtained for trials to extinction, al-

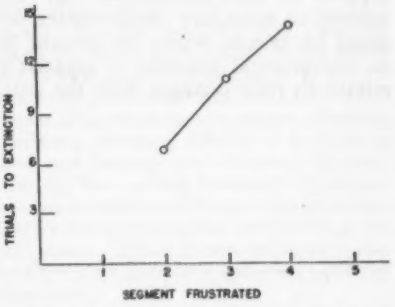


FIG. 1. Mean trials to extinction as a function of frustration at different distances from the goal

though the differences between the groups were not statistically significant. It is also apparent that the differential frustrations did not produce any significant differences in maze performance. The systematic verbal reports showed some tendency for general satisfaction with the experiment to be associated with frustration at a greater distance from the goal, but none of the obtained differences was statistically significant.

Discussion

The present experiment has demonstrated that the point at which frustration is introduced into an instrumental action sequence is a significant determinant of the strength of goal-directed behavior. It has been shown that human *S*s who have received equal practice and achieved the same criterion of proficiency in performing an instrumental action sequence exhibited differences in the number of attempts made to reach the goal, when they were frustrated at different distances from the goal. In general, these differences revealed a slightly negatively accelerated functional relationship between number of trials to extinction and number of instrumental acts completed during the goal-directed sequence. The present results had been predicted on the basis of Hull's goal gradient hypothesis and in general lend support to the assumption that the amount of secondary reinforcement accrued by stimuli which are present in an instrumental sequence is negatively related to their distance from the goal. Although the negatively accelerated gradient obtained is in keeping with curves previously obtained in goal gradient phenomena, it is likely that the nature of the functional relationship between resistance to extinction and distance from the goal at which frustration occurs may be expected to vary with changes in the experimental conditions involving modifications of the number of segments, drive level employed, or the amount of effort required to make a response.

Summary

Three groups of 13 human *S*s were trained to perform a series of 20 motor movements, which were divided into five discrete segments of four movements each. The *S*s in Group A were frustrated at Segment 2, Group B at Segment 3, and Group C at Segment 4.

The results indicate that the group which was frustrated closer to the goal engaged in the most trials to extinction, with number of trials to extinction decreasing progressively as frustration is introduced at further distances from the goal. Other measures of response strength failed to reveal significant differences.

An explanation of the phenomena obtained is proposed in terms of Hull's goal gradient hypothesis.

References

1. Adelman, H. M. Extinction of goal directed behavior as a function of frustration at various distances from the goal. Unpublished master's thesis, Wayne Univ., 1951.
2. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
3. Lambert, W. W., & Solomon, R. L. Extinction of a running response as a function of distance of block point from the goal. *J. comp. physiol. Psychol.*, 1952, **45**, 269–279.

(Received August 27, 1953)

Journal of Experimental Psychology
Vol. 47, No. 6, 1954

INTERPOLATION ACCURACY AS A FUNCTION OF VISUAL ANGLE BETWEEN SCALE MARKS

CHARLES A. BAKER

Aero Medical Laboratory, Wright Air Development Center

Interpolation accuracy is an important factor in the use of visual displays. Displays may be simple rulers or complex fuel gauges, mach meters, cathode ray tubes, or other similar devices. Grether and Williams (4), Carr and Garner (2), and Leyzorek (5) have studied interpolation accuracy as a function of the distance between the reference marks. All three studies indicate that the relative error (average error expressed as a percentage of the distance between marks) decreased with mark separation distance up to .5 in., and thereafter no decrease in errors resulted. Grether and Williams used circular scales, and Carr and Garner, linear scales. Leyzorek, however, used a series of concentric circles (range rings) separated by various distances. The present study is concerned with interpolation of this latter type of display.

Since Leyzorek's *S*s viewed the task at a distance of about 18 in., the visual angle subtended by the .5-in. range ring separation is 1° 36′. This visual angle value has subsequently been used to determine the optimal distance between range ring markers on cathode ray tube indicators. The implication is that the optimal range ring separation, expressed in inches on the scope face, varies with *S*'s viewing distance. However, none of the previous studies has actually varied the viewing distance with the range ring separation to test this hypothesis. The present paper is concerned with the effects of both range ring separation and viewing distance upon interpolation accuracy on a polar coordinate display.

METHOD

Apparatus.—The stimulus material was presented on 25 9 × 9-in. white cards. On each card was a circle 8 in. in diameter. Within the 8-in. circle on each card were placed 0, 1, 3, 7, or 15 equally spaced concentric circles. Thus the concentric circles (range rings) were separated by ¼, ½, 1, 2, or 4 in., respectively. Each of the five range ring separations appeared on five cards, making a total of 25 cards. A dot appeared in the center of each card. The lines used in outlining the circles were black and were .01 in. wide. Five targets or "pips" appeared on each of the 25 cards. The lines representing targets were ¼-in. arcs that were .025 in. wide. Different target locations were used for each of the five cards for a given range ring separation.

The 125 targets were distributed about equally in each quadrant of the display and in each ¼-in. interval from the center of the circle to the circumference. During the experiment *S*'s head movements were minimized by a chin rest. The card was placed in a stand that could be placed at distances of 10, 15, 20, 30, and 40 in. from *S*'s eyes. The plane of the cards was normal to the line of sight. The luminance of the cards was held constant at 17 ft.-L.

Experimental design.—The order of presentation of conditions was counterbalanced by the following procedure: The total of 25 cards was divided into five groups with five cards per group. Each group contained one card for each of the range ring separations. The counterbalancing of the card groups and the viewing distances was accomplished by the use of a 5 × 5 greco-latin square. The order in which a given *S* was presented the various range ring separations within any card group was counterbalanced by using a 5 × 5 latin square constructed for that *S*. Thus, sources of bias resulting from practice or learning effects and differential difficulty of the stimulus cards were balanced over all viewing distances. Also, any bias resulting from order of presentation was balanced over all range ring separations. One *S*'s data consisted of one row from the greco-latin square. Fifteen *S*s were employed so that each row of the greco-latin square was replicated three times.

Subjects.—The *S*s were 11 psychologists and 4 shop personnel, all males, employed by the Aero Medical Laboratory. The average age was 29

years, with a range from 21 to 33. All *S*s reported having 20/20 vision. They were not told the purpose of the experiment.

Procedure.—A range value of 1,000 yd. was represented by the separation between adjacent range rings regardless of the separation distance. Previous investigations (1, 3) indicate that this interval value for scales is superior to other values. The following instructions were given: "Here we have a series of simulated radarscopes with various numbers of range rings. The distance between range rings is equal to 1,000 yards. Your task is to estimate the position of the target between the range rings. In every case the inner adjacent ring is zero yards and the outer adjacent ring is 1,000 yards. Therefore, you will call out ranges varying from zero to 1,000 yards. Give your ranges to the nearest 10 yards. Be as accurate as possible and take as much time as you wish. Start with the target at the 12 o'clock position and call out the ranges as the targets appear in a clockwise direction."

Each *S* had a practice session of at least 15 min. on practice cards. Some required more time until satisfactory performance was achieved. The average time taken per *S* to call out 125 ranges was 100 min.

Results

The data were recorded in yards of range as called out by *S*s. These data were then converted into error scores. This was accomplished by obtaining the difference between the reported range and the actual range for each target. The direction of error was also noted. The range of each target to the nearest 10 yd. was determined with the use of a finely graduated rule by two *E*s independently, with redetermination until the *E*s concurred. The error scores in yards were then converted into absolute error scores in inches associated with the appropriate

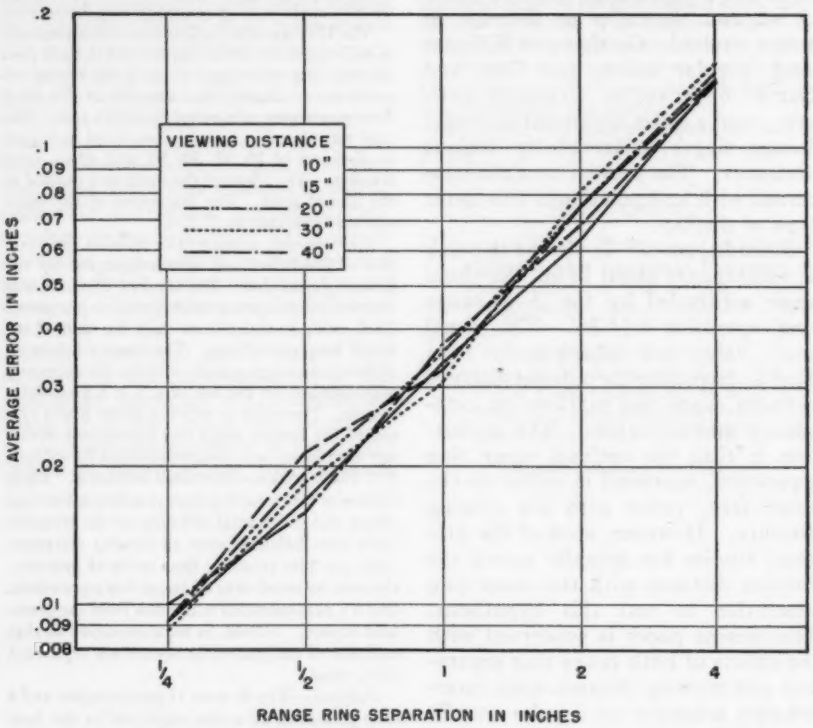


Fig. 1. Average error in inches as a function of range ring separation for each viewing distance

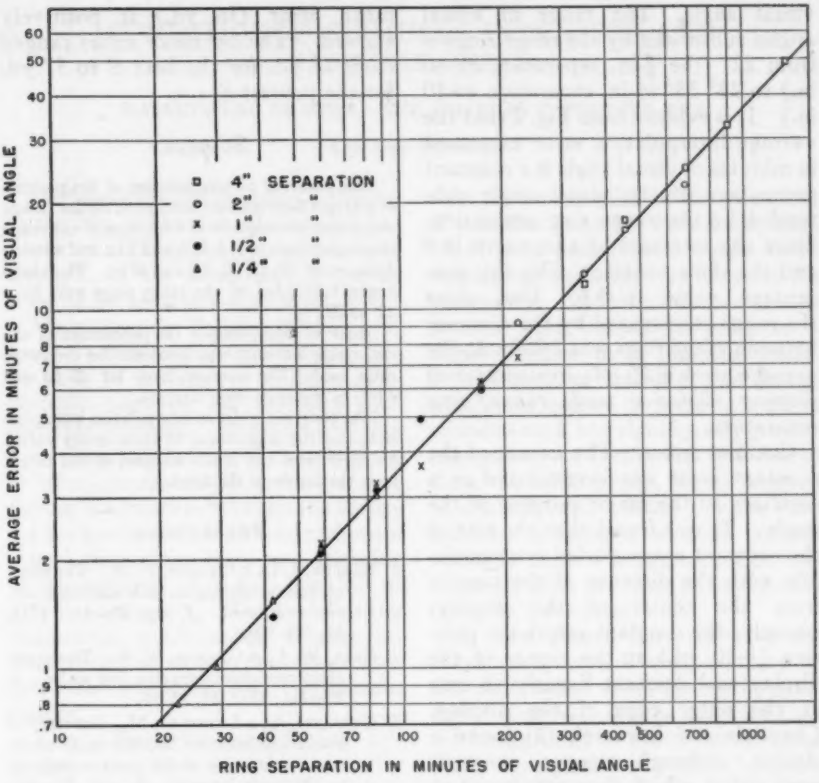


FIG. 2. Average interpolation error in minutes of visual angle as a function of the visual angle subtended by the various range ring separations at each viewing distance. The slope of the linear function is .036 and the intercept is 0.

range ring separation. These average error values are presented in Fig. 1 for all *Ss* averaged. The plot of the average error in inches against range ring separation is a linear increasing function for each viewing distance.

It is evident from Fig. 1 that viewing distance had no effect upon interpolation accuracy. The slope constant for all viewing distances averaged is .036. Leyzorek (**5**) did not vary viewing distance and reported that the average viewing distance was about 18 in., a distance *Ss* found convenient for the task. Leyzorek's finding and those of the present study agree quite well for range ring separations of .5 in. and above. However, Leyzorek found that for range ring separations below .5 in. the average error remained constant at about .02 in. Leyzorek used stimulus material that was of lower contrast and definition than the stimulus material used in the present study. This may account for the break in his curve at a range ring separation of .5 in.

Figure 2 is a plot of the average error against range ring separation, both expressed in terms of minutes of

visual angle. The range of visual angles subtended by the range rings is from 22′ (the ¼-in. separation at 40 in.) to 22° 38′ (4-in. separation at 10 in.) It is evident from Fig. 2 that the average interpolation error expressed in minutes of visual angle is a constant percentage of the visual angle subtended by the range ring separation. Since the intercept of this curve is 0 and the slope constant .036, this percentage value is 3.6. Also, since the range represented by the distance between range rings was 1,000 yd., the average error is 36 yd., irrespective of viewing distance and range ring separation.

Constant errors.—The nature of the constant error was investigated as a corollary to the major purpose of the study. It was found that the size of the constant errors varied systematically with the distance of the targets from the center of the display. Namely, the constant errors are positive (+20 yd.) at the center of the display and decrease linearly to zero at the outer edge of the display. Chapanis and Leyzorek (3) found a similar, although steeper, constant error curve, which they attributed to calibration errors. The present results suggest that this effect was not entirely due to such errors. Since Barber and Garner (1), who used rectangular coordinates (straight line grids), did not find these positive constant errors of interpolation, it is further suggested that these errors result from an illusion created by the curved range rings.

Subject variance.—The distribution of *S*s' mean errors about the grand mean error (36 yd.) is positively skewed. The *S*s' mean errors ranged from 25 yd. for the best *S* to 57 yd. for the poorest *S*.

Summary

The accuracy of interpolation of the position of a target between two concentric circles (range rings) was investigated as a function of the range ring separations of ¼, ½, 1, 2, and 4 in. and viewing distances of 10, 15, 20, 30, and 40 in. The visual angles subtended by the range rings were from 22′ to 22° 38′.

Error of interpolation (in percentage of the range ring interval) was invariant for the range rings used. The average error for all *S*s was 3.6% of the range ring interval.

The constant errors of interpolation were positive, and the magnitudes of these errors varied inversely with the radial distance of the target from the center of the display.

References

1. Barber, J. L., & Garner, W. R. The effect of scale numbering on scale-reading accuracy and speed. *J. exp. Psychol.*, 1951, **41**, 298–309.
2. Carr, W. J., & Garner, W. R. The maximum precision of reading fine scales. *J. Psychol.*, 1952, **34**, 85–94.
3. Chapanis, A., & Leyzorek, M. Accuracy of visual interpolation between scale markers as a function of the number assigned to the scale interval. *J. exp. Psychol.*, 1950, **40**, 655–667.
4. Grether, W. F., & Williams, A. C. Speed and accuracy of dial reading as a function of dial diameter and angular separation of scale divisions. In P. M. Fitts (Ed.), Psychological research on equipment design. *Aviat. Psychol. Res. Rep.* No. 19. Washington: U. S. Government Printing Office, 1947. Pp. 101–109.
5. Leyzorek, M. Accuracy of visual interpolation between circular scale markers as a function of the separation between markers. *J. exp. Psychol.*, 1949, **39**, 270–279.

(Received September 8, 1953)

Journal of Experimental Psychology
Vol. 47, No. 6, 1954

CONDITIONED STIMULUS INTENSITY AND RESPONSE SPEED

RAYMOND M. BRAGIEL[1] AND CHARLES C. PERKINS, JR.[2]

Kent State University

Hull has included stimulus intensity dynamism, a positive relationship between stimulus intensity and response strength, in the latest sets of postulates in his behavior system (3, 4).

More recently Perkins (6) has presented a differential conditioning hypothesis to explain the positive relationship between CS intensity and response strength. According to this hypothesis, extinction of the response to situation-minus-CS (which occurs during the interval between conditioning trials when nonreinforced responses occur) generalizes more to a stimulus of less intensity than to a stimulus of greater intensity. Thus, following differential conditioning, a more intense stimulus tends to produce greater response strength, i.e., dynamism results.

The differential conditioning hypothesis simplifies behavior theory by eliminating the *postulate* of dynamism and by increasing the specificity of the term "stimulus intensity." Thus, although it has been supported by a Skinner-box study (6) in which dynamism was obtained following differential training but not following nondifferential training, further experimentation is needed to determine the adequacy of this interpretation.[3]

In his formulation of the postulate of stimulus intensity dynamism, Hull (2, 3) gives considerable weight to an experiment performed by Hays in his laboratory. In this study two groups of rats were trained to jump to a single window in a jumping stand. At the limit of learning, one group, which was trained to jump to a white stimulus card, had significantly greater response speed (reciprocal of latency) than the other group, which jumped to a black card. Since a white card may be considered a more intense stimulus than a black one, this was taken by Hull (2, 3) as an instance of stimulus dynamism.

At first glance the results of the Hays study would seem contrary to the differential conditioning hypothesis. However, if it is assumed that the background in which the stimulus cards were placed was black,[4] Hays's results could have been due to the contrast between stimulus card and background. Thus, an application of the differential conditioning hypothesis to the development of orienting responses could be used to explain the results. Presumably, orienting toward the background is extinguished,

[1] Now at Girls' Industrial School, Beloit, Kansas.

[2] This study is a portion of the first author's M.A. thesis. The second author, who directed the thesis, is responsible for the preparation of this report.

[3] The differential conditioning hypothesis is not intended to explain a positive relationship between intensity of a stimulus with drive properties and response strength. For example, faster running to escape a strong shock than to escape a weak one cannot be satisfactorily derived from the differential conditioning hypothesis. However, such a result is analogous to a positive relationship between strength of the hunger drive and speed of running to food. Thus, faster running with a strong shock can be predicted from the difference in motivation and does not require a postulate of dynamism.

[4] The Hays study has not been fully described; details of the apparatus and procedure apparently have not been recorded and are no longer known. (Personal communication from Mr. John A. Antoinetti, 1953.)

and the extinction effects generalize to the stimulus card. If the stimulus card is the same color as the background (black card on black background), the effects of extinction will generalize to the card more than if they contrast (white card on black background). Thus orientation toward the white card on a black background would be more rapid than orientation toward the black card on a black background. Since appropriate orientation must occur before the response is made, the white-card (contrast) group would be expected to have greater response speed than the black-card group.

The present study consists of two experiments. In the first the background is black; one group of rats is trained to jump to a black card, the other to a white card. This serves as a repetition of what appear to be the essential conditions of the Hays experiment. The second experiment differs from the first only in that the background is white. This serves as a crucial test between absolute intensity and degree of contrast as determiners of response strength. It also serves as a check on the adequacy of the differential conditioning explanation of dynamism since this hypothesis clearly implies greater response strength (speed of response) for the contrast group.

Experiment I

Method

Apparatus.—A single (6 × 6 in.) window jumping stand was used. The movable (2 × 19 in.) jumping platform was the same height as the bottom of the stimulus window. Interchangeable black and white landing compartments were used. Automatic timing of the response was obtained by an electric timer in a circuit that was held closed by the weight of *S* while his center of gravity was on the 4 in. of the jumping platform nearest to the stimulus window.

The stimulus cards were 7 × 7-in. squares of 1/16-in. Plexiglas, painted black or white. The black paint was the same as that used on the background. Since the stand was constructed for use with either one or two windows, there was a 6 × 14-in. section inserted within the background wall. This removable section was clearly outlined to normal human vision. The stimulus cards were recessed ½ in. behind the surface of the background and thus were very clearly outlined to human vision even when they matched the color of the background. Air currents from a wave of the hand were sufficient to knock over the stimulus card.

The background was painted flat black. The apparatus was surrounded by a black cloth curtain which served as a one-way screen. Lighting for the jumping platform and stimulus cards was supplied by a 15-w. fluorescent light in a white metal reflector fastened to the top of the stand approximately 2 ft. above the jumping platform. The landing compartment was lighted by a 75-w. frosted glass bulb suspended approximately 4 ft. above the floor of the goal compartment. In all respects not specifically mentioned above, the apparatus was identical to that used by Ehrenfreund (1).

Subjects.—The *S*s were 12 experimentally naive albino rats. They were between 90 and 150 days old at the beginning of the experiment. The *S*s in the two groups were counterbalanced for age, strain, and sex. Four *S*s that had been used in another study were also run, but their results were discarded because their performance indicated marked transfer from previous training.

Preliminary training.—Gradual reduction of body weight to 87% of normal was effected by limited feedings at 24-hr. intervals. The *S*s were then placed on the jumping platform, which touched the open stimulus window. The platform was gradually moved back from the window in a regular fashion on subsequent trials. At the end of ten days and 90 trials, all rats were regularly jumping an 8-in. gap to the open window. On Days 11–14 the stimulus card was introduced. If *S*s refused to jump within 3 min., the stimulus card was removed and the gap shortened temporarily. On subsequent trials the gap was gradually increased to 8 in., and the stimulus card reintroduced. All *S*s were jumping regularly to the card by the end of Day 14.

Training period.—On each of Days 15–27 every *S* received ten trials to the stimulus card with an 8-in. gap. Throughout all training, every response of entering the landing compartment was reinforced by allowing *S* to eat (for 15 sec.) from a dish containing wet mash. The intertrial interval averaged about 3.5 min. For Group W–B a white landing compartment and white stimulus card were used. For Group B–B the landing compartment and stimulus card were

black. The *Ss* were run in squads of four consisting of two *Ss* in each group.

Results

The results during the training period are presented in the form of learning curves in Fig. 1. Each point in this figure was determined in the following manner: (*a*) Median response speed (reciprocal of latency) for each *S* on the last nine trials of the day was calculated. (The first trial was not used in this computation because of a warm-up effect. This is similar to Hull's treatment [3] of Hays's data.) (*b*) The median of these measures for all *Ss* in the group on the given day was determined. These medians are the points plotted.

Examination of Fig. 1 indicates learning during the first part of the training period, but no systematic changes in performance during the last five or six days. Also, after the first day there is greater response speed for Group W–B than Group B–B.

The significance of the difference between the two groups was tested by the Mann-Whitney *U* test (5). The *U* based on individual *Ss*' medians for the last nine trials was 3.5. The *U* for median response speed on the last 27 trials (not including first trials of the day) was 4.0. A *U* of 4 is significant at the 1.3% level of confidence (single-wing hypothesis). The results of Exp. I confirm Hays's findings in that, with black background, *Ss* learned to respond more rapidly to the white card.

Experiment II

Method

Apparatus.—The entire background of the apparatus was painted a flat white matching the white stimulus card. In other respects the apparatus was the same as that used in Exp. I.

Subjects.—The *Ss* were 20 naive albino rats between 90 and 120 days old at the beginning of preliminary training.

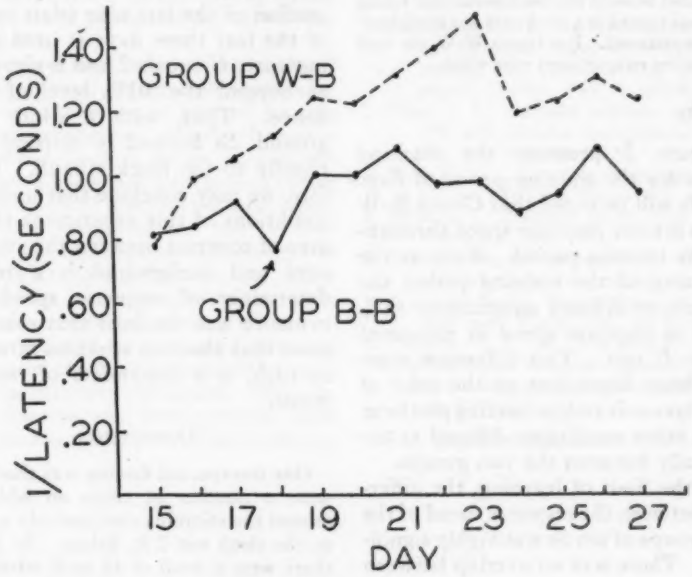


Fig. 1. Learning curves of Exp. I

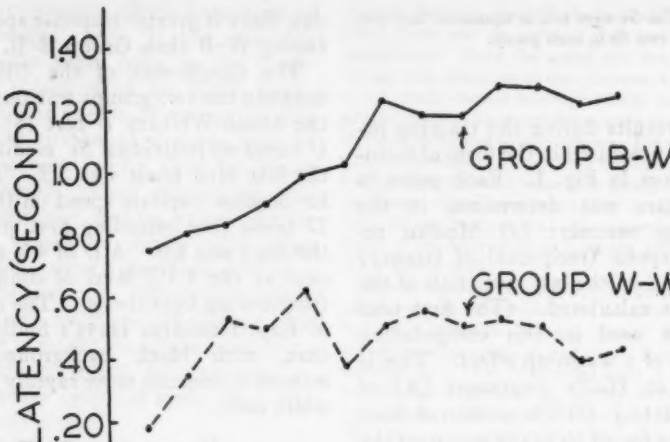


FIG. 2. Learning curves for Exp. II

Procedure.—The procedure was identical to Exp. I. Only the color of the background differentiated between the two experiments. Group B–W was trained to a black card and black landing compartment. For Group W–W the card and landing compartment were white.

Results

Figure 2 presents the learning curves for the training period of Exp. II. It will be noted that Group B–W shows greater response speed throughout the training period. Even at the beginning of the training period the two groups differed significantly (1% level) in response speed as measured by the *U* test. This difference must have been dependent on the color of stimulus cards and/or landing platform as no other conditions differed systematically between the two groups.

At the limit of learning, the difference between the response speed of the two groups of ten *S*s was highly significant. There was no overlap between the groups in median response speed of individual *S*s on the last nine trials of learning (i.e., $U = 0$). When the median of the last nine trials on each of the last three days is used as the measure, *U* equals 2 and is significant at beyond the .01% level of confidence. Thus, with a white background, *S*s learned to respond more rapidly to the black stimuli. Therefore, we may conclude that under the conditions of this experiment the degree of contrast between the stimulus card and background is a reliable determiner of response speed. No evidence was obtained that even suggests that absolute stimulus intensity, as such, is a determiner of response speed.

DISCUSSION

One unexpected finding was that there were a number of trials on which *S*s missed the stimulus card entirely and fell to the cloth net 2 ft. below. In Exp. I there were a total of 49 such misses distributed among the six *S*s in Group B–B,

and three misses among two *Ss* in Group W–B. In Exp. II there were 102 misses distributed among all ten *Ss* in Group W–W, and only a total of three misses among two *Ss* in Group B–W. The greater frequency of misses when the stimulus card and background were the same color suggests that jumping to the card generalized to the background, and that there was differential conditioning of the jumping response. This interpretation is supported by a tendency for response speed to decrease following misses (although *S* always jumped within the 3 min. allowed).

However, there is also evidence which indicates that differential training of orienting responses was involved. In Exp. II only two *Ss* (both in Group W–W) missed the stimulus card after Day 23. Therefore, it appears unlikely that the effects of misses, which should extinguish with further training, could be entirely responsible for the consistently longer latencies of Group W–W at the end of the training period. Thus, it seems that differential conditioning of orienting responses rather than of jumping responses was primarily responsible for the results on the last day of training in Exp. II.

The major implication of the results of Exp. II is that if the principle of stimulus intensity dynamism is applied to this type of experiment, stimulus intensity must be defined in terms of contrast between stimulus and background rather than in terms of absolute physical intensity. The results should not be considered as contradicting the principle of stimulus intensity dynamism, but rather as defining more precisely just what determines stimulus intensity.

The results of our study also support a differential conditioning interpretation of stimulus intensity dynamism. Since we have demonstrated that contrast must have been responsible for Hays's results, her data cannot be taken as evidence against the differential conditioning hypothesis.

Summary

Two groups of six rats were trained to jump in a single stimulus window jumping stand with a black background. The procedures were identical except that for Group W–B the stimulus card and landing compartment were white, while for Group B–B the card and landing compartment were black. At the limit of learning, response speed was significantly (1.3% level) greater for Group W–B.

Two further groups of ten *Ss* each were run in a second experiment. For both groups the background was white. The stimulus card and landing compartment were black for one (B–W), and white for the other (W–W). Group B–W had significantly greater response speed at the limit of learning.

On the basis of these and other findings, it was concluded that: (*a*) stimulus intensity is better defined in terms of contrast with background than in terms of absolute physical intensity; (*b*) these results and those of Hays are in agreement with a differential conditioning interpretation of stimulus intensity dynamism.

References

1. Ehrenfreund, D. An experimental test of the continuity theory of discrimination learning with pattern vision. *J. comp. physiol. Psychol.*, 1948, **41**, 408–422.
2. Hull, C. L. Stimulus intensity dynamism (*V*) and stimulus generalization. *Psychol. Rev.*, 1949, **56**, 67–76.
3. Hull, C. L. *Essentials of behavior.* New Haven: Yale Univ. Press, 1951.
4. Hull, C. L. *A behavior system.* New Haven: Yale Univ. Press, 1952.
5. Mann, H. B., & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. math. Statist.*, 1947, **18**, 50–60.
6. Perkins, C. C., Jr. The relation between conditioned stimulus intensity and response strength. *J. exp. Psychol.*, 1953, **46**, 225–231.

(Received September 14, 1953)

Journal of Experimental Psychology
Vol. 47, No. 6, 1954

DISTRIBUTED PRACTICE AND ROTE LEARNING IN CONCEPT FORMATION [1]

JACK RICHARDSON AND BRUCE O. BERGUM [2]

Northwestern University

Psychologists interested in concept formation appear to agree that the process can be described as the learning or the discovery of relationships. The lack of agreement beyond this point is reflected in the tasks used to demonstrate the process. These range from the sorting of objects of various forms, colors, and materials into groups (4), to the association of nonsense syllables with pictorial objects presented on a memory drum (5). There is also little agreement as to whether or not there is a commonality of the processes underlying concept learning and rote learning. Since the processes underlying learning must be inferred from performance, it seems that the only way to demonstrate the commonality of the processes is to determine the variables and their effects on each type of learning. If the variables and their effects are the same, then there is no reason to infer different processes.

Some degree of distribution of practice has facilitated the learning of a wide variety of lower-order tasks (7), but the effect on the more complex tasks has not been clearly demonstrated. Studies of mental mazes (2), problem solving (3), and puzzle solving (1) indicate that massed practice is superior to distributed practice, but the intertrial interval used in the distributed condition was never less than 24 hr. Oseas and Underwood (6), using 6-, 15-, 30-, and 60-sec. intertrial intervals, found that distribution tended to facilitate simple concept learning, but the differences were not statistically significant.

The first experiment in the present report is essentially a reproduction of the Oseas and Underwood study under conditions which were considered optimum for facilitation of learning by distribution. The concepts were presented at a faster rate, and different dimensions were used as distracters in an attempt to vary the difficulty of the lists. The results of this experiment were inconclusive, and a second experiment was designed to test hypotheses suggested by the first.

EXPERIMENT I

Method

Materials.—The stimuli were simple geometric figures that varied along three dimensions (form, size, and shade). Three levels of each dimension were used. The figures were either circles, triangles, or squares; small, medium, or large ($\frac{1}{4}$, $\frac{3}{4}$, or $\frac{1}{2}$ in. in height); and white, gray, or black. All possible combinations of the three levels of the three dimensions permit 27 different figures. These were drawn with ink on white vellum tape. Single letters of the alphabet, selected so that there were no obvious associations between the stimulus figures and the letters, were typed on the tape and used as responses to designate the concepts.

Lists.—Three different lists were constructed, each with the same figures as stimuli and the same letters as responses. In each of the lists the concepts were systematically selected using two of the dimensions as relevant to the concepts and one dimension as irrelevant to the concepts. Each list contained three different examples of nine concepts; the only difference in the three lists was the irrelevant dimension. For example, in one list form was used as the distracter, and size and shade were relevant to the concepts.

[1] The authors wish to express their gratitude to Dr. B. J. Underwood, who assisted with the study.

[2] Now at Training Research Laboratory, Department of Psychology, University of Illinois.

In this list a small gray figure had the same response associated with it regardless of whether it was a circle, a triangle, or a square. Other letters were associated with different combinations of size and shade, i.e., small and white, small and black, medium and white, etc. One list had form as the distracter, one list had size as the distracter, and one had shade as the distracter.

The use of three levels of each dimension permitted three examples of each concept since only one dimension was not relevant to the concepts. Six different orders of the concepts were used in each list. Each order consisted of one illustration of each of the nine concepts. The position of the concept example within an order was randomly assigned, and the same example of any particular concept did not appear in two successive orders. Each order was separated from the next by a strip of masking tape; thus, each presentation of the nine concepts constituted a trial, and the position of an illustration of a given concept changed randomly from one trial to the next.

Subjects and conditions.—A table of random numbers was used to assign 120 college students to one of six conditions. The only differences in the conditions under which each group of 20 *S*s learned were the lists and the intertrial interval. Each list had a different dimension as the distracter and each was learned by massed (4-sec. intertrial interval) and by distributed (30-sec. intertrial interval) practice. With I, II, and III representing the groups that learned the lists using form, size, and shade, respectively, as the distracter, and with M and D representing massed and distributed practice, the six conditions were I–M, I–D, II–M, II–D, III–M, and III–D.

Each *S* received instructions about the definition of a concept as used in this experiment and about the procedure. The lists were presented by the usual paired-associate method at a 2:2-sec. rate; i.e., the stimulus was presented for 2 sec. and then the stimulus and response appeared together for 2 sec. The *S*s in the distributed-practice conditions named colors during all intertrial intervals. No practice was given to *S*s, and the appropriate experimental list was presented to each *S* for 25 trials regardless of the amount of learning.

Results

The mean number of correct responses given on the 25 trials by each group is presented in Table 1. A simple analysis of variance, using the within-groups variance as the error term, gives *F*'s of 10.51, 2.21, and 2.16 for the effects of the distracter, the distribution of practice, and the interaction between the distracter and the distribution, respectively. The differences due to the distracter are significant at the 1% level, while the differences due to distribution and to interaction are not statistically significant.

TABLE 1

MEAN NUMBER OF CORRECT RESPONSES

Condition	I (Form)	II (Size)	III (Shade)	σ_m
Massed	53.90	70.45	72.60	7.40
Distributed	50.15	75.10	98.65	

The finding that distribution of practice is not a significant variable in this experiment seems to be misleading when the trends are analyzed further. Although the interaction between the distracter and the distribution is not significant, inspection of Table 1 shows that distribution of practice has a greater tendency to facilitate learning in the group that gave the most correct responses. The concepts with form as the distracter were the most difficult, and the massed group is slightly superior to the distributed group. The distributed group that learned concepts with size as the distracter is slightly superior to the massed group, and, with shade as the distracter, the distributed group is significantly superior ($t = 2.49$; $p < .05$).

Concept learning, as used in this experiment, may logically be divided into three stages, which must follow in sequence in order for the concepts to be learned: (*a*) the identification of the dimensions and the levels of each used in the illustrations of the concepts; (*b*) the discovery of the dimensions relevant to the concepts; and (*c*) the association of the concepts with the letters assigned to them. It is probable that all of these stages occur simultaneously when all concepts are considered, but it seems that the learning must take place in this sequence for any particular concept.

The first stage of learning is the same for all three lists. If we assume that the response is associated with equal ease with any combination of two dimensions (Stage 3), then the difference in the difficulty of the lists is due to the difference in difficulty of discovering the dimensions relevant to the concepts (Stage 2). If these three stages are discrete processes, then it is possible that distribution of practice will have different effects upon the different stages and produce the results obtained in this experiment.

The second experiment was designed to isolate the three stages experimentally, to determine the relative contribution of each stage to the total task, and to provide further evidence as to the discreteness of the processes involved in concept formation and rote learning.

Experiment II

Method

Material and subjects.—The list used in this experiment had shade as the distracter and was the same as the corresponding list in the preceding experiment. The presentation of the list was the same except that all learning was by massed practice to a criterion of one perfect trial.

The *Ss* (120 college students) were systematically assigned to one of four groups. The first four to appear at the laboratory were assigned to Cond. I through IV, the second four to Cond. IV through I, etc. throughout the study.

Four sets of instructions were used, one for each group. The instructions for Cond. IV were those used in Exp. I, and the other instructions deviated from it only as demanded by the different conditions. Rest periods of 2 min. were interpolated at those points in conditions other than Cond. I corresponding to the intervals during which *Ss* in Cond. I were instructed about the procedure for the succeeding step. The *Ss* were not actively employed during these intervals.

Conditions.—In each condition the concepts were learned to a criterion of one perfect trial, but the four conditions were divided into parts corresponding to various combinations of the three stages described previously. Part 1 consisted of identifying the dimensions and the levels of each; Part 2 consisted of discovering the dimensions relevant to the concepts; and Part 3 of associating the correct letter name with the concept. The *Ss* in Cond. I learned the three parts separately and were required to attain one perfect trial on each part before proceeding to the next. The learning in Cond. II was in two sections; the first section consisted of Part 1, and the second section of a combination of Parts 2 and 3. Thus, after identifying the dimensions and levels, the *Ss* were instructed to call out the correct letter as soon as possible. Condition III was also in two sections; the first section was a combination of Parts 1 and 2, and the second section consisted of Part 3 only. In Cond. IV the *Ss* learned all three parts together. This condition was comparable to the learning in Exp. I. To summarize: Cond. I learning consisted of the three parts separately; Cond. II of Part 1 and then Parts 2 and 3 together; Cond. III of Parts 1 and 2 together and then Part 3; and Cond. IV of learning all three parts as one task.

In Cond. I, II, or III, when Parts 1 or 2 appeared as separate tasks in the procedure, *Ss* were instructed to call out the common name of the levels, i.e., large black circle, etc. The stimulus figures for these discrete steps were presented for 4 sec. although the response letter did not appear. If *S* called out the correct levels, a buzzer was sounded during the last 2 sec. of presentation.

Results

Trials to learn.—The mean number of trials to learn each part and the combinations of parts for each condition is presented in Table 2. A comparison of the total trials required to learn all three parts for each of the conditions gives an F of .03. The comparison of the trials required to learn Parts 1 and 2 in Cond. I and the trials required to learn Parts 1 and 2 in Cond. III gives an F of 1.19. Since neither of these is significant, they indicate that there is no differential

TABLE 2

Mean Number of Trials to Learn

Group	Part					
	1	2	3	1&2	2&3	1&2&3
I	2.43	3.47	29.93	6.16	33.66	36.09
II	2.03				37.30	39.33
III			32.97	4.50		37.47
IV						36.30

effect of learning by the "part" or by the "whole" method.

The data from Cond. I were analyzed to determine the approximate contribution of each part to the total trials to learn. The resulting percentages were 7, 10, and 83 for Parts 1, 2, and 3, respectively. In Cond. III, Parts 1 and 2 combined contributed 12% and Part 3 contributed 88% of the total trials to learn. The major proportion of the total task is clearly one of simple rote memorization.

Correlations between parts.—In Cond. I the correlations between the trials required to learn Part 1 and Part 2, Part 1 and Part 3, and Part 2 and Part 3 are −.12, −.20, and .25, respectively. Since the *SD* of Part 1 is only 1.31 and the *SD* of Part 2 is .249, the low correlations may be due to the restricted range of the scores on Parts 1 and 2.

Discussion

The main findings of this study were: (*a*) changing the dimension used as the distracter had significant effects on learning, and (*b*) the large component of rote learning in the task used to demonstrate concept formation tended to obscure the concept learning.

In Exp. I the group that learned the list with form as the distracter gave significantly fewer correct responses than the groups that learned the lists with size or shade as the distracter. Logically, there is no reason for selecting any two dimensions as relevant to the concepts in preference to any other two. Similarity has been suggested as one of the variables of concept learning (**8**), and it seems that increasing the similarity of the levels of the dimension used as the distracter should facilitate learning, while increasing the similarity of the levels of the relevant dimensions should inhibit learning.

In order to check this hypothesis, 34 *S*s were asked to rank the dimensions according to the similarity of the levels within the dimension. The figures were drawn on white vellum cloth to the same specifications as those used in the experiment. The three strips of cloth each had three figures drawn on it and the three figures on each strip varied in only one dimension. Table 3 shows the number of times that each dimension was ranked in each place with 1 representing the most similar. It is evident that the levels of form are the least similar, with little difference between the similarity of the levels of size and the levels of shape. As a further check, the 27 stimulus figures were drawn on cards, and 82 *S*s were asked to sort the cards into nine piles in three different ways.[3] This may be done by using one of the dimensions as irrelevant to each sorting. The results agreed well with the previous ranking of similarity of the levels; form was least often used as irrelevant first in the sorting, with little difference in size and shape. The different degrees of learning of the lists in this study can be attributed to the difference in similarity of the levels of the dimensions used as distracters. The most difficult list had form as the distracter and the levels of form were the least similar.

TABLE 3

Rank Order of the Similarity of the Levels of the Dimensions

Rank	Dimension		
	Form	Size	Shade
1	1	17	16
2	4	13	17
3	29	4	1

If it is assumed that distribution of practice has a different effect upon concept learning as opposed to rote learning, then the differential effects of distribution of practice resulting from the variation of the difficulty of the lists can be explained. The least difficult list had the largest component of rote learning and was facilitated most by distribution. The most difficult list had the largest

[3] Robert A. Gardner suggested this method of ranking similarity and collected the data on the card-sorting task.

component of concept learning and was facilitated least by distribution. Oseas and Underwood (6), using form as the distracter, found that distributed practice tended to facilitate learning. This appears to contradict the results of the present study, where the comparable list was learned slightly better by massed practice. However, *S*s in the Oseas and Underwood study learned to a criterion of two successive perfect trials, while in the present study *S*s gave a mean of 4.02 correct responses on the last trial. Learning to a higher criterion would increase the time spent on rote learning and thus increase the probability that distributed practice will facilitate learning.

Over 80% of the task in Exp. II was rote learning, and analysis of the data of Exp. I trial by trial indicates that the most difficult list could hardly have required more than 50% of the 25 trials for concept learning. The task used in these experiments is obviously not appropriate for demonstrating concept learning. Further experimentation must be done with a task in which rote learning has been either eliminated or demonstrated to be of minor importance.

Summary

The first experiment in this study used simple geometric figures varying in form, size, and shade. Three lists, differing only in the dimension designated irrelevant to the concepts, were learned by massed and by distributed practice. The list that had form as the distracter was significantly more difficult than the lists that had size or shade as the distracter. Distribution of practice was not a significant variable, but there was some tendency for distribution to facilitate the learning of the least difficult list. Logically the learning seemed to consist of three stages: (*a*) identifying the dimensions and the levels of each, (*b*) discovering which dimensions are relevant to the concepts, and (*c*) associating the concepts with the designated responses (letters of the alphabet).

The second experiment was designed to isolate the three stages of learning and to determine the relative contribution of each stage to the total trials to learn. There were no significant differences in the trials required to learn the stages separately or simultaneously. The major proportion of the task used in this study was simple rote learning since the third stage took over 80% of the total trials required to learn the complete task.

References

1. Cook, T. W. Massed and distributed practice in puzzle solving. *Psychol. Rev.*, 1934, **41**, 330–355.
2. Cook, T. W. Factors in massed and distributed practice. *J. exp. Psychol.*, 1944, **34**, 325–334.
3. Ericksen, S. C. Variability of attack in massed and distributed practice. *J. exp. Psychol.*, 1942, **31**, 339–345.
4. Goldstein, K., & Scheerer, M. Abstract and concrete behavior. An experimental study with special tests. *Psychol. Monogr.*, 1941, **53**, No. 2 (Whole No. 239).
5. Heidbreder, E. The attainment of concepts: II. The problem. *J. gen. Psychol.*, 1946, **35**, 191–223.
6. Oseas, L., & Underwood, B. J. Studies of distributed practice: V. Learning and retention of concepts. *J. exp. Psychol.*, 1952, **43**, 143–148.
7. Underwood, B. J. *Experimental psychology*. New York: Appleton-Century-Crofts, 1949.
8. Underwood, B. J. An orientation for research on thinking. *Psychol. Rev.*, 1952, **59**, 209–220.

(Received October 1, 1953)

VISUAL–MOTOR PERFORMANCE AS A FUNCTION OF SHORT–DURATION AMBIENT TEMPERATURE [1]

WARREN H. TEICHNER AND ROBERT F. WEHRKAMP

Quartermaster Research and Development Center

Viteles and Smith (4) observed that pursuit performance, as well as performance on other types of tasks, shows a loss when the still air temperature is raised above the comfort range. Mackworth (2) has confirmed this for visual-motor performance, at least with respect to tracking with a heavily weighted control. It seems reasonable that decrements in visual-motor performance should also appear below the comfort range and that some type of curve with a single maximum can be used to express the relationship. The present study was designed to test this hypothesis.

In the study by Viteles and Smith (4), six *Ss* performed a variety of tasks over a 4-hr. test period. Similarly, Mackworth's (2) ten *Ss* alternated performance with a heavily weighted and a lightly weighted control. There is the possibility, therefore, that at least part of the decrements observed in these studies was attributable to effects produced by the alternation of the tasks and by the effects of prolonged effort. A second purpose of the present study, therefore, was to provide a control for these possible effects. This was done by using only one task and having *Ss* perform for only a short period of time.

Method

Apparatus.—The apparatus was a Koerth-type pursuit rotor with a pressureless (hinged) stylus. An electric clock graduated in .01 sec. was used to measure time on target. An electronic relay system governed the starting and stopping of the rotor and thus controlled the length of each trial and the time between trials.

Subjects.—The *Ss* were 30 Army enlisted men. Each *S*, clothed in an Army fatigue uniform, sat quietly in a room with a constant temperature for 20 min. prior to being tested. Then while standing he was given 15 trials on the pursuit rotor; each trial lasted 20 sec., with a 10-sec. interval between trials. This procedure was repeated once daily for five successive days at the same time every day. The highest score of each *S* was posted daily to provide *Ss* with knowledge of their performance and resulted in overt signs of competition for "best score."

The *Ss* were arranged randomly into five temperature groups, viz., 50°, 55°, 70°, 85°, and 100° F., with ten *Ss* in the 70° F. group and five in each of the other groups. Two *Ss* had to be withdrawn from the 70° F. group because of illness, thus leaving eight *Ss* in this group. In addition, the data of the entire 50° F. group had to be discarded when it was discovered that the cooling system had not held this temperature reasonably constant. Thus, four groups were available for analysis, three with five *Ss* each and one, the 70° F. group, with eight.

Results and Discussion

Previous investigators (1, 3) who have studied the effects of stress (particularly fatigue) on psychomotor performance have experienced considerable difficulty in obtaining decrements in spite of imposing extremely rigorous conditions upon their *Ss*. This difficulty might be the result of (*a*) use of averages over a whole performance period, thereby losing the possibility of greater sensitivity of the earlier or later part of performance of the task and/or (*b*) failure to consider adequately the large learning factor involved in most psychomotor tasks and a possible compensatory gain in performance, which might mask any possible decrements. To test the first hypothesis, the 15 trials obtained daily

[1] This study was reported to the 1953 meeting of the American Psychological Association.

were divided into blocks of 5 trials each, comprising the first, second, and third part of the daily performance. Mean time on target was calculated for each such block of trials, and these means were used as raw scores. In order to evaluate the learning factor, these raw scores were considered for each of the five experimental days. Figure 1 shows the effect of the temperature on the learning of the task. each part of this graph presenting a different third of the performance period. Inspection of this figure indicates that acquisition on the rotor pursuit proceeded quite rapidly from day to day for all thirds of the performance period, and further, that performance was affected by the temperature used. It also suggests that the degree to which measures were influenced differed according to which portion of the performance period is considered. Thus, for the second third, it appears that response strength increased most rapidly in order for the 70°, 85°, 55°, and 100° F. groups. The same trends and the same differences between groups also appear in the last third, and here they appear slightly more clearly than they do on the second third. On the other hand, Fig. 1 shows that with the exception of the 100° F. group, all groups performed in the same way in the first third of the task.

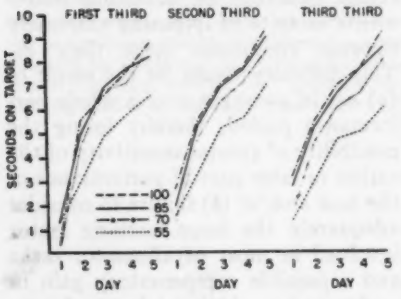


FIG. 1. Daily acquisition of each third of performance under four ambient temperature conditions

TABLE 1

SUMMARY OF ANALYSIS OF VARIANCE OF TIME ON TARGET IN BLOCKS OF FIVE TRIALS

Source of Variation	*df*	Mean Square	*F*
Temperature	3	12.16	71.53*
Days	4	74.10	437.87*
Periods	2	.35	
T × D	12	.08	
T × P	6	.18	
D × P	8	.30	
Error	24	.17	
Total	59		

* $p < .01$.

An analysis of variance of the data of Fig. 1 is summarized in Table 1. The temperature effects and the differences between days were significant at well beyond the 1% level of confidence. It may be concluded, therefore, that the differences among the groups of Fig. 1 are real and not the result of random selection. It may also be stated that the daily gains in performance are real gains. On the other hand, the differences between thirds of performance periods did not prove to be significant; *F* in this case fell short of the 5% level of confidence. The fact that both the Periods × Days and the Periods × Temperatures interactions also were not significant lends additional support to the hypothesis that there are no real differences between periods. Further, separate analysis of the data into blocks of three trials each, rather than five, again confirmed the null hypothesis with respect to differences between portions of the practice period.

In order to assess the kind of relationship holding between temperature and visual-motor performance as tested here, the data of Fig. 1 were replotted as shown in Fig. 2. This

latter figure presents the strength of performance as a function of ambient temperature for each of the five days separately. This is done once for the second and once for the third of the daily period. This figure suggests that the function reaches a maximum around 70° F. and falls off from this value in both directions. The figure also suggests that performance is poorer in temperatures lower than 70° F. than for those higher than this value. This is indicated by the fact that a temperature of 15° less than 70° F. (i.e., 55° F.) produced poorer performance than one 15° greater than 70° F. (i.e., 85°). These suggestions are repeated with very few deviations throughout the five days and for both of the periods.

To determine the general effect of the temperatures over all aspects of the daily performance, all three periods were averaged arithmetically. These resulting means and *SD*'s are presented in Table 2. Inspection of this table indicates that these average effects were similar to those shown in Fig. 2 for the second and third periods. The most general indication of the effect of temperature may be seen by inspection of the grand means for the temperatures shown in Table 2.

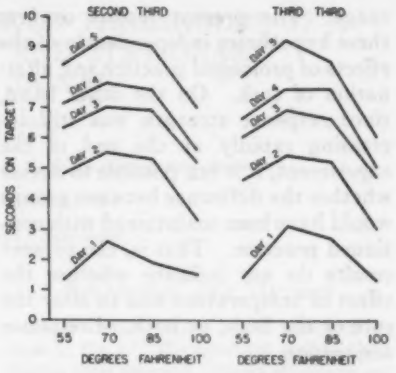


FIG. 2. Mean daily performance of the second and third portions of performance as a function of short-duration ambient temperature

Again the data indicate that the curve falls in both directions from a maximum of 70° F. and that it falls more precipitously toward the low temperature side.

Even though the present *S*s performed for only a short time upon a single task, the results confirm those of Viteles and Smith (4) and Mackworth (2) in that response strength was found to be less the higher the ambient temperature above the comfort range. In addition, the data indicate that response strength is less when the temperature is below the comfort

TABLE 2

MEAN TIME ON TARGET BY DAYS

Day	Temperature							
	55° ($N = 5$)		70° ($N = 8$)		85° ($N = 5$)		100° ($N = 5$)	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
1	1.16	.94	2.48	1.68	1.93	1.54	1.26	1.37
2	4.64	1.24	5.49	2.39	5.18	2.54	3.15	2.22
3	6.37	1.84	6.82	3.04	6.80	2.28	4.92	2.92
4	6.93	2.10	7.62	2.65	7.63	2.33	5.51	2.99
5	8.33	1.86	9.12	2.78	8.59	2.65	6.52	2.80
Total	5.49	2.97	6.31	3.24	6.03	3.27	4.27	3.15

range. The present results confirm these hypotheses independently of the effects of prolonged practice and alternation of task. On the other hand, since response strength was still increasing rapidly at the end of the experiment, it is not possible to decide whether the difference between groups would have been maintained with continued practice. That is, the present results do not indicate whether the effect of temperature was to alter the rate or the limit, or both, of response acquisition.

Summary

Time-on-target measures on the pursuit rotor were obtained from four subject groups which practiced for short periods on five successive days under ambient temperatures of 55°, 70°, 85°, and 100° F., respectively. Performance was found to be poorer at temperatures above and below 70° F. This result confirms and extends previous studies of the effects of ambient temperature. In addition, the data suggest the possibility that the temperature function falls off more rapidly with temperatures lower than the maximum than it does for those that are higher.

References

1. Finan, J. L., Finan, S. C., & Hartson, L. D. A review of representative tests used for quantitative measurements of behavior-decrement under conditions related to aircraft flight. *USAF Wright Air Development Center, Tech. Rep.*, July 1949, No. 5830.
2. Mackworth, N. H. Researches on the measurement of human performance. London: His Majesty's Stationery Office, 1950. (Med. Res. Counc., Spec. Rep. Ser. No. 268.)
3. Teichner, W. H. Recent studies of simple reaction time. *Psychol. Bull.*, 1954, **51**, 128–149.
4. Viteles, H. S., & Smith, K. R. An experimental investigation of the effects of change in atmospheric conditions and noise upon performance. *Heat., Pip., Air Condit.*, 1946, **18**, 107–112.

(Received September 25, 1953)

Journal of Experimental Psychology
Vol. 47, No. 6, 1954

COMPARISON OF THE MASKED THRESHOLDS OF A SIMULATED MOVING AND STATIONARY AUDITORY SIGNAL [1]

R. C. WILCOTT

University of Southern California and U. S. Navy Electronics Laboratory

AND R. S. GALES

U. S. Navy Electronics Laboratory

The work described herein has been designed to determine whether or not the masked threshold of a signal moving in perceived azimuth is lower than that of a fixed azimuth. Descriptions are given of four separate experiments which were designed to determine the effects of simulated movement on the masked threshold under several conditions. It is well known that when an *O* is listening with earphones to a signal on a binaural system, the signal can be made to appear lateralized by introducing a time delay into one channel. The degree of lateralization, or distance the signal appears to be off center, is, within limits, a function of the amount of time delay. On this basis an auditory signal can be made to appear to move by introducing a continuously changing amount of time delay into one channel. By means of an electrical delay line coupled with a multiple-unit rotary switch, linear stepwise changes in interaural time delay that elicited a clear perception of auditory movement were produced.

[1] This report covers work on the joint University of Southern California–U. S. Navy Electronics Laboratory Project No. 1. It is part of a series of investigations under Contract No. NObsr-57435 between the Bureau of Ships, Navy Department, and the University of Southern California, in cooperation with the U. S. Navy Electronics Laboratory. The experimental work was carried on at the U. S. Navy Electronics Laboratory.

Apparatus and Method

The basic components of the apparatus are shown in Fig. 1.[2] The signal was produced by a thermal noise generator with an attenuator and filter for controlling the intensity level and width of the noise band. Turning on and off of the signal and its presentation time were controlled by a cuer that operated an electronic switch. Beyond the switch the signal circuit divided into two channels. The linear stepwise change in interaural delay was produced in one of the signal channels by a 60-step rotary switch coupled to a 60-unit variable electrical delay line (1). Each unit of the delay line consisted of a carefully constructed low pass filter which produced 9.0 microseconds of time delay. Thus, the range of the variable delay line was from 0 to 540 microseconds in increments of 9.0 microseconds. The rotary switch was turned by a constant speed motor operating through a variable speed changer. In addition to the variable delay, a constant 270-microsecond delay line, which consisted of 30 low pass filter units and which could be switched into the other signal channel, was used. It was switched into the circuit to produce the "frontal movement" described below.

The primary consideration in the selection of the signal was to secure a noise band which provided a clearly perceived simulation of movement. Since it has been demonstrated in the literature on localization that interaural time differences are very effective in determining the localization of low frequency components, a noise band containing frequencies between 75 and 850 cps was selected. It was found that the perception of movement was easily recognized with this band.

An uncorrelated masking noise, which was produced by two thermal noise generators, each separately feeding into the two signal channels, was used. The resulting masking noise at the

[2] The authors are indebted to C. E. Cunningham of the U. S. Navy Electronics Laboratory for his advice and help in the construction of the equipment.

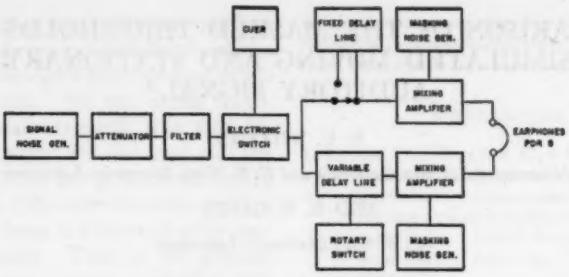


FIG. 1. Block diagram of apparatus

two ears has zero interaural correlation because the phase relationships of the noises produced by the two separate generators are continuously changing in a random order. The reasoning leading to the use of such a masking noise follows:

Recent experiments have shown that the level of masking of low frequency tones by noise is a function of the correlation of the masking noise (2). Preliminary observations definitely showed that the level of masking of the noise band selected as the signal in this study was a function of the interaural correlation of the masking noise. With a correlated (both ears in phase) masking noise produced by using one noise generator feeding the two binaural channels and with maximum time delay of the signal in one channel, the resulting threshold was found to be about 8 to 10 db below the threshold obtained with no time delay of the signal in either channel. Under this condition, the movement in azimuth produced by the variable delay line was sufficient to cause the signal to go from a level of clear audibility when localized at the side to inaudibility near the center. To minimize this effect and to secure a measure of the masked threshold of a simulated moving signal devoid, as much as possible, from this complicating factor, the uncorrelated masking noise was used. Even with an uncorrelated masking noise, it was found that the threshold of the signal is still somewhat a function of its localization. With this condition and when maximum time delay is placed in one channel, the threshold is 2 to 4 db *above* the level obtained with zero time delay.

The intensity of the masking noise was set at a constant level of .1 v. as measured across the earphones and, at this level, it produced about 35 db of masking of the signal when it was localized at the center position. It consisted of the full range of frequency components produced by the noise generators as modified by the PDR–8 earphones.

The *O*s whose ears were matched within 5 db for frequencies below 1500 cps were selected. During threshold measurement, *O* was seated in an acoustically insulated listening booth, with *E* occupying an adjoining room.

Two locations of movement were selected for study. These are shown in Fig. 2. The lateral movement appears to begin at about the position shown in Fig. 2, approximately 70° off center, and proceeds to the center position. This was produced by sweeping through the maximum delay to zero delay in 9.0-microsecond increments (540 to 0 microseconds). The frontal movement appears to begin at about the position in Fig. 2, approximately 35° off center, and proceeds past the center position to a point an equal distance to the other side. It was produced by setting up a constant delay in one of the signal channels of 270 microseconds and then introducing the variable delay in the other channel, sweeping from the maximum of 540 to 0 microseconds. When the variable delay reaches the 270-microsecond point, the delay in both channels is equalized, and thus the signal is heard as being at the center position.

EXPERIMENT I

The purpose of Exp. I was to determine the effect of simulated movement on the masked threshold with movement at the lateral and frontal positions in the phenomenal field and with two different velocities. The two velocities selected were: (*a*) a perceptually "slow" rate of movement, with the signal being moved over the complete distance permitted by the equipment in 6 sec., and (*b*) a perceptually "fast" rate of movement, with the complete distance being covered in 2 sec. Simulated movement at both the lateral and frontal positions and

at the two rates was readily apparent to all *O*s.

The standard method of limits, using only a descending series with the signal being attenuated in 1-db steps, was used. The fact that the threshold of the signal was dependent on its localization posed the problem of selecting a test method that would provide comparable thresholds for both fixed and moving cases. The method used in Exp. I required that the signal with simulated movement be heard by *O* during the entire time in which it was moving. Without this constraint, the threshold established might be a function of the particular localization of the signal at any given instant in time. To insure that *O* heard the signal during the entire period in which it was presented, a small light was coupled to the electronic gate so that it went on and off with the signal. The *O* was instructed to make sure he heard the signal go on and off synchronously with this light. Each *O* was trained until he reported five consecutive threshold values with a range of 3 db or less.

Results

The results indicate that under these conditions a simulated moving signal does not elicit a lower masked threshold than is elicited by a stationary signal. This conclusion is definitely indicated by the data given in the tables. Each set of measurements was obtained for each *O* during the same experimental period. A counterbalanced design was used.

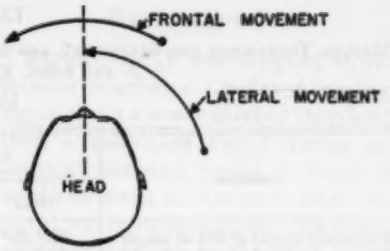


FIG. 2. Mean location and approximate distance of simulated movement in the phenomenal field

The mean threshold values for lateral movement are presented in Table 1. When a comparison is made between the threshold for a stationary signal at the lateral position and a moving signal when *O* is instructed to make sure he hears the signal go on and off, the differences are found to be less than 1 db at both the 6-sec. and 2-sec. rates of movement. Furthermore, the differences between the threshold for a stationary signal at the center position, and the threshold for the moving signal when *O* is instructed to attend to the cessation of the signal whether he heard it go on or not, are within 1 db at both the 6-sec. and 2-sec. rates.

TABLE 1

MASKED THRESHOLDS FOR STATIONARY AND MOVING SIGNALS AT THE LATERAL POSITION WITH 2- AND 6-SEC. RATES OF MOVEMENT

($N = 6$)

Condition	6-Sec. Rate		2-Sec. Rate	
	Mean	Range	Mean	Range
A. Moving signal with *O*s listening for onset and cessation	22.1 db*	18.7–23.1 db	25.1 db	22.1–28.3 db
B. Stationary signal at the lateral position	21.8	18.2–23.5	24.4	23.1–28.8
Difference (A–B)	0.3		0.7	
C. Moving signal with *O*s listening for cessation only	25.7	21.9–27.5	25.7	24.9–29.8
D. Stationary signal at the center position	25.2	21.1–27.0	25.1	21.1–29.1
Difference (C–D)	0.5		0.6	

* Threshold values are attenuation below an arbitrary reference level.

TABLE 2

MASKED THRESHOLDS FOR STATIONARY AND MOVING SIGNALS AT THE FRONTAL POSITION WITH 2- AND 6-SEC. RATES OF MOVEMENT

($N = 5$)

Condition	6-Sec. Rate		2-Sec. Rate	
	Mean	Range	Mean	Range
Stationary signal at left of center	23.9 db*	21.3–27.9 db	24.6 db	23.6–28.3 db
Stationary signal at right of center	24.5	21.1–27.8	24.2	22.0–28.2
Moving signal with *Os* listening for onset and cessation	23.9	21.3–27.8	25.6	23.5–29.6
Stationary signal at center position	25.7	23.6–29.0	25.9	21.8–28.2

* Threshold values are attenuation below an arbitrary reference level.

These comparisons show that the threshold obtained with a moving signal is a function of the localization of the signal at any given instant in time, and that movement per se has a small but practically insignificant effect on the threshold level. This conclusion is supported by the secondary observations of the *Os*. They all reported that when they were instructed to make sure they heard the moving signal go on and off, the signal seemed to increase in loudness as it moved toward the center position.

The same pattern of results was obtained with frontal movement. These are shown in Table 2. When the threshold was determined with the signal localized to the left of center or localized to the right of center, about the same level is obtained as the threshold for the moving signal when *O* is instructed to make sure he hears it go on and off. Any differences present are within 1 db. As would be expected, these three threshold levels tend to be slightly higher than the threshold obtained with the signal localized at the center position.

EXPERIMENT II

In Exp. II a faster rate of movement was used with an added effect. The signal was moved through the entire distance in 1 sec. and, instead of making just one excursion across the phenomenal field during each presentation, the signal was moved over and back four times.

This was accomplished by turning the rotary switch at a continuous rate of 1 rps. This produced the effect of moving the signal across in 1 sec. and then returning it to the starting point instantaneously, etc. The presentation time of each signal was 4 sec., and therefore, with a 1-sec. rate of movement, the signal made four excursions across the phenomenal field during the time of each presentation. With this procedure the perception of movement appeared more pronounced than with just one cycle of movement at the 1-sec. rate. Threshold measurement was the same as that employed in Exp. I except that *Os* were not instructed to specifically attend to either the onset or cessation of the signal. They were instructed to merely judge whether or not they heard the signal at any time during the 4-sec. presentation period.

Results

Mean threshold values are shown in Table 3. It is clear that the masked thresholds for a simulated moving signal with a 1-sec. rate of movement, and at both the lateral and frontal positions, were not significantly different from those obtained with the signal stationary at the center position.

TABLE 3

MASKED THRESHOLDS FOR STATIONARY AND MOVING SIGNALS AT LATERAL AND FRONTAL POSITIONS WITH A 1-SEC. RATE OF MOVEMENT

($N = 5$)

Condition	1-Sec. Rate	
	Mean	Range
Lateral movement	20.7 db*	18.5–23.9 db
Stationary signal at center position	21.8	18.9–25.3
Frontal movement	21.0	18.3–24.8

* Threshold values are attenuation below an arbitrary reference level.

EXPERIMENT III

Experiment III was designed to determine whether or not simulated movement would have any effect on the masked threshold when *O* was actively engaged in producing and controlling the movement himself.

This was accomplished by fastening a hand crank to the shaft of the rotary switch. During the presentation of the signal, *O* was instructed to make one complete revolution of the switch at a rate which maximized the perception of movement for him. Before beginning the threshold measurements, each *O* was given a period during which he was instructed to determine the rate of movement, both frontal and lateral, that appeared to elicit the most pronounced perception of movement. After he was satisfied with a particular rate of movement, he was instructed to try to maintain this rate constant during the threshold specification period. The presentation time of the signal was 6 sec. and was immediately preceded by the cue light. The *O*s were instructed to begin turning the switch as soon as the cue light went off.

Results

Mean threshold values are shown in Table 4. It is apparent that simulated movement produced and controlled by *O* has no appreciable effect on the masked threshold.

EXPERIMENT IV

Experiment IV was designed to determine whether a simulated moving signal elicits a lower masked threshold than a stationary signal during an extended listening period in which *O* would be given no cue as to when the signal was to be presented. This procedure was set up on the basis of the subjective impression that a moving signal has greater "attention value" than a stationary signal. In other words, the supposition was that a moving signal would attract the attention of *O* more readily and at a smaller signal-to-noise ratio than a stationary signal.

The *O*s were required to listen for a 2-hr. listening period that was divided into four parts and separated by 10-min. rest periods. To prevent *O* from knowing when a threshold determination was to be made, the method of limits was used, with only ascending series being given. In addition, the temporal order of the threshold determination runs was systematically randomized in such a manner that each threshold determination for either a stationary or moving signal was preceded by a former threshold determination, either a stationary or moving signal, and by time intervals ranging from 2 to 9 min. All *O*s reported that under this condition they had no cue as to when the signal was to be presented other than the fact that they were suddenly aware of its presence. Only the frontal movement was used with this procedure.

TABLE 4

MASKED THRESHOLDS FOR STATIONARY AND MOVING SIGNALS AT LATERAL AND FRONTAL POSITIONS WITH RATE OF MOVEMENT CONTROLLED BY *O*

($N = 5$)

Condition	Mean	Range
Lateral movement	22.5 db*	19.7–26.1 db
Stationary signal at center position	22.2	20.1–25.8
Frontal movement	21.9	20.2–23.8

* Threshold values are attenuation below an arbitrary reference level.

TABLE 5

MASKED THRESHOLDS FOR STATIONARY AND MOVING SIGNALS AT A FRONTAL POSITION DURING AN EXTENDED LISTENING PERIOD ($N=5$)

Condition	Mean	Range
Stationary signal	19.3 db*	18.6–20.4 db
Moving signal	18.7	18.0–20.0

* Threshold values are attenuation below an arbitrary reference level.

Results

Mean threshold values are shown in Table 5. Again, it is apparent that a moving signal does not elicit a lower masked threshold than a stationary signal. These results, therefore, do not support the subjective impression, reported by these and other *O*s, that a moving signal has greater "attention value" than a stationary signal.

DISCUSSION

The results of these experiments do not support the hypothesis that the masked threshold of a signal moving in perceived azimuth is lower than that in a fixed azimuth. In the interpretation of this conclusion it must be remembered that the auditory movement employed was a simulated movement produced by a linear stepwise change in interaural delay, and it is thus not the same as real movement. During real movement, varying interaural intensity differences would be present, and under the condition of movement away from or toward the head, a doppler effect would be present. However, as the movement simulated by this procedure is readily perceived and elicits the illusion of real azimuthal movement in the phenomenal field, it is reasonable to assume that the conclusion formulated above would apply to real movement in azimuth, disregarding any gains in actual signal strength resulting from the movement.

SUMMARY

The masked thresholds of a simulated moving signal produced by a linear stepwise change in interaural time delay, and with movement at two different locations in the phenomenal field, were compared with the thresholds obtained with a stationary signal. A noise band extending from 75 to 850 cps was used as the signal and was presented against an uncorrelated background noise, and masked thresholds were made under the following conditions: (*a*) with a signal moving at a perceptually slow and fast rate; (*b*) with a perceptually fast-moving signal making four excursions across the phenomenal field during each presentation period; (*c*) with the signal movement controlled by *O*; and (*d*) during an extended listening period in which *O* was given no cue as to when the signal was to be presented. The results of the threshold comparisons with these different conditions indicate that the masked threshold of a moving signal is a function only of its localization at any given instant in time, and that a moving signal does not elicit a lower masked threshold than does a stationary signal.

REFERENCES

1. GALES, R. S., & KLUMPP, R. Use of electrical delay lines in research on binaural hearing. *J. Audio-engng Soc.*, in press.
*2. HIRSH, I. J. The influence of interaural phase on interaural summation and inhibition, *J. acoust. Soc. Amer.*, 1948, 2, 536–544.

(Received September 25, 1953)

Journal of Experimental Psychology
Vol. 47, No. 6, 1954

ERRORS IN TRANSFER FOLLOWING LEARNING WITH UNDERSTANDING: FURTHER STUDIES WITH KATONA'S CARD-TRICK EXPERIMENTS[1]

ERNEST R. HILGARD, ROBERT D. EDGREN,[2] AND ROBERT P. IRVINE

Stanford University

Students who have learned simple card tricks with understanding are able to transfer their learning to new tasks more successfully than those who have learned by rote memorization. This is the major finding of an earlier study with high school students (1), supporting Katona's similar results with college students (2). The most puzzling feature of the results in the earlier study was the large number of errors the "understanding" *Ss* made. Although they did better on new tasks than the "memorization" *Ss*, one-third of the understanding group made errors on the simplest transfer task, and three-fourths made errors on the most difficult task. The arrangements of the experiment permitted perfect scores, based solely on the problem-solving methods taught in the course of experimentation. These prevalent errors by the understanding *Ss* set the problem for the present investigation.

In order to get at the source of errors made by *Ss* as they attempted to transfer from a learned task to a novel one, we designed five variations of the "understanding" method of training. We conjectured that the kind of diagramming or other process used to teach the problem to *S* might affect transfer by providing a different kind of generalization in each case. Learning cannot be classified satisfactorily as either rote memorizing or learning by understanding; not only do the two processes overlap, but neither process is unitary. We know that *Ss* can learn meaningful material more readily than nonsense, so that some "understanding" enters into "rote" memorization. Similarly, a memorized formula permits generalized learning ("understanding") within the limits of the data that can be fitted appropriately into the formula. We hoped that our different methods of training might help us to discover some of the "dimensions" along which degrees of understanding might ultimately be scaled. While the differences produced by our five variations proved disappointingly small, the results throw some light on the question with which we began: Why are so many errors made in transferring understanding to new problems?

METHOD

The experiments were conducted individually with 150 high school students from communities in the neighborhood of Stanford University. The *E* arranged for space at each school, and *Ss* came for the experimental sessions on each of two successive days. The *Ss* were assigned at random to each of the five experimental groups, following an order determined from a table of random numbers. To assure an equal number of each sex in each group, assignments were made independent of sex until the sex quota of any one group was completed. After that, an *S* who would have been assigned to a group whose sex quota had already been filled was given instead the immediately following assignment. A departure from these rules occurred at the very begin-

[1] This experiment is one in a series of studies of problem solving being done under Project NR 150-104 and supported by Contract Nonr 225 (02) between Stanford University and the Office of Naval Research. Work on the contract is under the general direction of Dr. Donald W. Taylor. Permission is granted to the United States Government for reproduction, translation, publication, use, and disposal of this article in whole or in part.

[2] Now at Mills College.

ning of the experiment, when only three methods were used instead of five, but the rules were followed in completing assignments to these first three groups as well as in filling the remaining three groups. Data were collected between August 1952 and July 1953, ending with 30 *Ss* in each group, equally divided between men and women.

The procedures, except for the variations in methods of teaching understanding, were those described in detail in the earlier report (1). Day 1 began with a foredemonstration by *E*, followed by training on two tasks (Task 1 and Task 2). Day 2 began with repetition of the foredemonstration, then a retest of Tasks 1 and 2, followed by tests for transfer on four additional tasks (Tasks 3, 4, 5, and 6). Each task required *S* to arrange playing cards in a fixed order. For example, in order to meet the requirements of Task 1, *S* had to arrange eight cards in order so that, dealing the cards from the pack first one face up on the table and then one to the bottom of the pack, the cards would appear face up in the order R–B–R–B–R–B–R–B. To do this he had to begin with the cards in the pack in the order R–R–B–R–R–B–B–B. The tricks are all of this general type, the chief complication being that some cards (those played to the bottom of the pack) are used more than once before these cards are dealt to the table. The order in which *S* arranged the cards was scored simply as correct or incorrect, although *E* recorded the exact order of incorrectly arranged cards so that he could make a later study of errors.

The six tasks are listed below, with their solutions. Additional details appear in the earlier paper. The order of tasks is that used on the second day, when tests of retention and of transfer were made.

Tasks for Original Training and for Testing Retention

Task 1: 4 red, 4 black alternating; skip one. (*Solution:* R–R–B–R–R–B–B–B)

Task 2: 8 spades in numerical order; skip one. (*Solution:* A–5–2–7–3–6–4–8)

Task Testing Transfer by Transposition from Task 1

Task 3: 4 even, 4 odd alternating; skip one. (*Solution:* E–E–O–E–E–O–O–O)

Tasks Testing Transfer by Problem Solving

Task 4: 3 red, 3 black alternating; skip one. (*Solution:* R–B–B–B–R–R)

Task 5: 4 red, 4 black alternating; skip two. (*Solution:* R–R–B–B–B–B–R–R)

Task 6: 10 spades in numerical order: skip two. (*Solution:* A–10–8–2–5–7–3–9–6–4)

The five methods of teaching understanding were as follows:

Method 1: Katona diagram.—The method is that described as the understanding method in the previous paper (1), and follows Katona (2, pp. 263–264). In it *S* diagrams the task by writing in horizontal rows what happens each time through the pack of cards as one card is dealt face up on the table and one to the bottom of the pack. The cards dealt face up are named in the diagram, while the cards placed on the bottom are indicated by a question mark. The requirements of the trick determine how the question marks eventually get assigned numbers or colors, so that a final summary of the diagram determines the correct arrangement of the cards in the pack.

Method 2: Horizontal row.—Because some *Ss* found it difficult to follow the several rows of the Katona diagram, we devised a modification in which all the information would be given in a single horizontal row. First, *S* wrote down the numbers from one to eight (or whatever the number of cards to be used in the trick). Then for the experiment in which the object was to arrange by colors, he used a red and blue pencil to indicate the red and black of the cards. In the experiment requiring alternately red and black cards to be played from the deck, skipping one between, *S* began by making a red mark under his numeral one, then a blue mark under three, red under five, blue under seven, thus indicating the color of the card required in each of these positions. Then he must return to the start, knowing that all numerals with colored marks have been played. Skipping number eight, he makes his next red mark at number two; then he skips the space at number four and places a blue mark at six, the next open space. Continuing in this way, he ends with a diagram with the positions of red and black cards all marked according to the requirement of the trick. For the trick requiring a succession of numbers to turn up, the same method is used, except that numerals are written down instead of red and blue marks. The two possible advantages from the method were: (*a*) colored marks directly related to the colors on the cards might prevent errors due to confusion between letters R and B; (*b*) the single row diagram might prevent errors made in the several-row diagram proposed by Katona.

Method 3: Vertical column.—Because *S* must finally arrange the cards in a pile from top to bottom, it occurred to us that further errors might be eliminated if the diagram were made as a vertical column instead of a horizontal row. Hence Method 3 was exactly the same as Method 2 except that the numbers representing the card positions were written in a column instead of in a

row, and the marks were made with the red and blue pencil beside these vertically arranged positions. The main error possibly avoided in this method is that of picking up the cards in reverse order, as sometimes occurs with the horizontal method. There is also a "naturalness" about recognizing what it means to place a card on the bottom of the pile, and using it again as the trick proceeds.

Method 4: Slips of paper.—In this method *S* uses numbered slips of paper to represent the cards in their original order, as distinguished from the row of numerals in Method 2 and the column of numerals in Method 3. By performing the trick with the blank pieces of paper, the requirements of the trick are easily met. All that *S* needs to do is to label each piece as it is played on the table red or black, if that is what the trick calls for, or A, 2, 3, etc., if a trick calls for such a progression. Because each slip of paper now bears a number indicating its original order in the pack, and a designation as to what kind of card it is, *S* can arrange the actual cards according to the scheme worked out with the blank papers. This method has the possible advantage that a symbolic step is eliminated because actual slips of paper are manipulated rather than symbolized orders in a diagram. It also appears to be widely generalizable to tricks of this type, because all *S* has to understand is the requirements of the trick. Then he performs the trick as called for, and does exactly the same sort of labeling for every trick.

Method 5: Working backward.—Taking advantage of what Piaget has called the "reversibility of thought" (5), we have taken for granted that *S* could understand that an orderly sequence in a forward direction implies an orderly sequence in the reverse direction. The conceptual difficulty common to all these card tricks is that some cards are reused several times in the process of dealing from the top to the bottom of the deck. The simplest way to overcome this difficulty is to begin with the last card played on the table, then to pick up the second to the last card played, and so on through the deck, observing the condition that a card played to the bottom of the deck in the forward order must be played from the bottom to the top of the deck in the reverse order. This method, while requiring a reversal of procedure not called for in any of the others, is actually the most direct. It is done with the playing cards themselves, without any intermediaries in the form of diagrams or substituted slips of paper.

We introduced the qualitative differences among the methods in the hope that they might result in differential errors. In the future, we would hope to find some way to quantify the differences in method, but at this exploratory stage we felt it necessary to have more qualitative information.

Results

Level of performance following each method of training.—Differences of success among the groups trained by the various methods should have appeared in the scores of the tasks of the second day. The percentages of success of *S*s trained by each method are given for each of the tasks in Table 1.

The over-all results lead to the same conclusions as those of the earlier study: Even though the methods used

TABLE 1

Percentage of *S*s Making Perfect Scores in Tests of Retention, Transposition, and Problem-Solving Transfer

(N = 30 for each method of training)

Method Used in Training	Percentage of *S*s Making Perfect Scores					
	Retention		Transposition	Problem-solving Transfer		
	Task 1	Task 2*	Task 3	Task 4	Task 5**	Task 6
1. Katona method	90	97	83	77	63	60
2. Horizontal row	80	77	63	77	23	37
3. Vertical column	67	87	63	70	30	60
4. Slips of paper	70	87	63	73	60	57
5. Working backwards	63	80	80	87	33	50
All methods (N = 150)	74	85	71	77	42	53

* Not a "pure" test of retention because of relearning of Task 1 before testing on Task 2.
** Only Task 5 shows significant differences among the five methods; chi square = 17.47, $.01 > p > .001$.

TABLE 2

ANALYSIS OF ERRORS IN MOST DIFFICULT TASK (TASK 5)

Nature of Error	Number of *Ss* Making Stated Error (N = 30 for each method)					
	Method 1 (Katona)	Method 2 (Horizontal)	Method 3 (Vertical)	Method 4 (Slips)	Method 5 (Backwards)	All Methods
1. Confusion based on requirement to skip two cards instead of one	10	20	12	9	12	63
2. Errors in technique attributable to method	1	0	5	6	12	24
3. Reverted to memorization	0	3	4	1	1	9
4. Red and black reversed	2	2	3	0	2	9
5. Still working when time called	2	1	0	1	0	4
Total errors	15	26	24	17	27	109
Number of *Ss* making these errors	11	23	21	12	20	87

here were all "understanding" methods, a substantial fraction of *Ss* made errors, especially on the transfer tasks.

There are suggestions of differential success according to method, but statistical tests of differences show a departure from chance only for Task 5, the most difficult transfer task. Task 5 is the first task in which *S* is required to skip two cards instead of one as he performs the trick. This was found to be the most difficult transfer problem in the earlier study. The increase in difficulty of Task 5 over Task 4 for the total 150 *Ss* is statistically significant using a chi square that takes into account the correlation between successes on the tasks (**4**, p. 206). The chi square is 37.55, yielding a p less than .001. When *S* is faced by this most severe test, the differences in method of training show up.

The departures from chance among the five groups on Task 5 that show up in the over-all test of significance ($.01 > p > .001$) are obviously produced by the high success of *Ss* trained by Methods 1 and 4, as compared with *Ss* trained by Methods 2, 3, and 5. The preferred methods are those described as the Katona method and the slips of paper method. These are superior, for Task 5, over the other three methods (horizontal row, vertical column, and working backwards). These differences provide an occasion for the study of the errors made by *Ss* trained with each method.

Analysis of errors on the most difficult transfer task.—The attempt to solve Task 5 following Task 4 proved more difficult for *Ss* of every group. While the transfer was made somewhat more successfully by those using the Katona method and the slips of paper method, all groups made enough errors for an analysis of the errors to be promising. The errors tabulated in Table 2 were in most cases noted at the time the task was performed, although in some cases it was necessary to infer the nature of the error from the erroneous order of cards appearing in the record.

By far, the most common type of error (63 of 109 errors) had to do with the problem of skipping two cards between those dealt on the table. This may be interpreted as in part a perseverative error, because all the

earlier tricks had called for the skipping of one card. There were 13 *Ss* who skipped one card throughout, despite the instructions to skip two. Some genuine confusion resulted among those who tried to follow the instructions, especially when they had dealt all but two cards in the deck. To skip two at this stage meant of course to deal the upper card to the bottom, then the new upper card to the bottom, thus returning the two cards to the position they had been in before starting this last double skip. Many *Ss* made an error at this point; it was the most common single error.

Next in order came errors differentially attributable to some features of the particular method.

Those *Ss* using the two horizontal diagram methods (Methods 1 and 2) were almost free of errors attributable specifically to the method, although the errors in skipping two places might be attributed in part to method. These errors occurred very commonly at the end of a line in the diagram, *S* commonly "losing his place" as he counted the remaining open spaces. Apparently the question marks in the Katona diagram were more effective "counters" than the unmarked serial positions of the horizontal method, for fewer errors of this kind were made by those using the Katona method.

The errors made in the vertical method did not differ greatly from those in the horizontal method. Place skipping in the diagrams accounted for most of the errors. The failure to find any advantage for the vertical arrangement over the horizontal one shows that caution is needed before making an inference from the "logic" of the arrangement to its psychological consequences; we had conjectured that the vertical arrangement would be superior because it "fitted" better the cards arranged in a pack.

The first three methods all belong to one family of diagramming methods. The advantage appears to lie with the Katona diagram over the other two new types of diagram.

The remaining two methods, slips of paper and working backwards, each gave rise to new types of error. An error unique to the slips of paper method was that of faulty labeling of the slips. Another error was a failure to arrange the cards in the order indicated by the slips. Thus the slips of paper served rather well as an intermediate "crutch," but they were by no means infallible.

The working backwards method led, for some undetermined reason, to the predominant error of skipping one card (instead of two) all the way. Half of the *Ss* who made errors in this group (10 of 20) did this, while the same error was made by only 3 of the 120 *Ss* working by other methods. Instructions were either misunderstood or misinterpreted; *S* placed the cards to be picked up in pairs on the table, in the order RR–BB–RR, as though instructions called for playing out the cards in pairs instead of playing two cards from the top to the bottom of the pack before dealing out a single card. One observation is pertinent. The working backwards method was very easily acquired and *S* proceeded with great confidence, even in error. Part of the misperception of instructions may have resulted from the very overconfidence engendered by the method; i.e., *S* plunged ahead, doing something by two's, without being clear as to just what was required.

The tendency to revert to a method of rote memorization appears when the "understanding" methods fail. Thus the methods of horizontal rows and vertical columns, which proved among the most difficult, account for seven

of the nine *S*s who still reverted to memorization at this stage of the experiment.

The conclusion appears to be justified that there are some differences in the error distributions by the different methods, although the numbers of errors dealt with are small, and sampling fluctuations are to be expected. Because the differences are largely based upon the arrangement of "counters" and "place-holders," it is a plausible conjecture that perceptual features of the kinds of diagramming used, or of card arrangements on the table, affected the result. The differences were not according to prior guesses, so that the perceptual features that are important are not the ones picked out in designing the experiment, e.g., the intimacy between the color of marks used for recording and the color of cards, and the arrangement in a single serial row or column.

A logical point of some importance that does not show up in the quantitative material is the greater ease of checking the result when the Katona method is used as against any of the other methods. The diagram is spread out in such a fashion that the original steps are retraceable. The method of working backwards is weakest in this respect, for once the cards are picked up there is no record whatever as to what went on. The slips of paper method allows little check also, although minor errors, such as slips sticking together, are more easily remedied than in the working backwards method. The prevalence of careless errors (such as reversing red and black) shows that it is important for the teaching of understanding to devise methods that are not only understandable in themselves, but are open for review and check, through retracing the steps, in order to correct mechanical and other errors that creep in even though *S* fully knows what he is doing.

Nature of the improvement between Tasks 5 and 6.—Task 5 was the first in which the task called for skipping two cards instead of one, and as the foregoing analysis showed, more errors were attributable to this requirement than to any other. The results in Table 1 show that there was a general improvement in score between Task 5 and Task 6 (except for the scores by Methods 1 and 4, which were the most successful on Task 5). Examination of this improvement may suggest something of what was learned in the experience with Task 5, for Task 6 also required skipping two cards.

First of all, is the improvement statistically significant? For the 150 *S*s, the 53% successful on Task 6 as compared with the 42% successful on Task 5 represent an increase near the .01 level (chi square = 6.4 for 1 *df*, allowing for the correlation of successes between tasks). The greatest percentage gain is for *S*s using the vertical columns method, in which 21 *S*s made errors on Task 5 and only 12 made errors on Task 6. Percentage gains were made also by *S*s using the horizontal diagram and working backwards methods, while *S*s using the other two methods showed no improvement.

In Table 3 the errors on Task 6 are compared with those on Task 5 for the three groups that improved between the two tasks. If we consider simply the decrease in total errors of a given kind between the two tasks, the greatest reduction (10 errors) is in the errors due to confusion over skipping two cards instead of one. Apparently some *S*s profited by the experience on Task 5, as we would expect. The next greatest reduction (7 errors) was the elimination of red-black reversals, but this error was not possible in Task 6,

TABLE 3

NATURE OF IMPROVEMENT BETWEEN TASK 5 AND TASK 6
(Methods 2, 3, and 5 only)

Nature of Error	Number of *Ss* Making Stated Error ($N = 90$)		Change in Errors between Tasks
	Task 5	Task 6	
1. Confusion based on requirement to skip two cards instead of one	44	34[1]	−10
2. Errors attributable to method	17	16	− 1
3. Reverted to memorization	8	4	− 4
4. Red and black reversed	7	0	− 7
5. Still working when time called	1	0	− 1
Number of errors	77	54	−23
Number of *Ss* making these errors	64	46	−18

which called for a numerical sequence of cards of one suit. Reverting to memorization, which led to errors in Task 5, was abandoned by some but not by all of the *Ss* who tried it. The errors attributable to the specific techniques persisted, or cropped up for other *Ss*. Of the 17 who made errors of this kind on Task 5, 9 made similar errors on Task 6, but 7 new *Ss* now made corresponding errors.

The general impression persists that carelessness, partly imposed by working with time limits, results in some kinds of error that appear or disappear almost by chance, with the exception that the total number of errors decreases when the task is easier, or after *S* has had more experience. The particular kind of error that shows up is difficult to account for in terms of lack of basic understanding of relationships inherent in the task; the kind of error appears to be determined rather by casual circumstances of method and generalized attitudes of caution or carelessness.

Sex differences.—Because we used an equal number of male and female students in each condition, and assigned them at random, we are free to compare scores according to sex. In order to obtain a measure of individual differences in performance, the six tasks of Day 2 were treated as a six-item test, each item of which could be passed or failed. Hence the possible scores ranged from 0 to 6. Because of the clustering of scores near the high end of the scale, computation of means and standard errors seemed inappropriate. Hence the scores were dichotomized, those *Ss* with scores between 0 and 4 falling in the lower group, those with scores of 5 and 6 in the higher group. Sex differences revealed by such scores are presented in Table 4. While the sex differences are small, and negligible in some comparisons, all comparisons favor the male *Ss*, and the total result is significant (chi square = 11.1, $p = .001$).

This incidental finding of male superiority among high school students in these problem-solving tasks is con-

TABLE 4

SEX DIFFERENCES ACCORDING TO METHOD OF TRAINING

Method of Training	Scores Based on Number of Tasks Solved Correctly on Day 2			
	Ss Making Lower Scores (0 − 4)		*Ss* Making Higher Scores (5 − 6)	
	Male	Female	Male	Female
1. Katona method	4	7	11	8
2. Horizontal row	11	12	4	3
3. Vertical column	9	11	6	4
4. Slips of paper	5	13	10	2
5. Working backwards	6	12	9	3
Total	35	55	40	20

sistent with the findings for problem solving and reasoning among college students as reported by other investigators (3, 6).

Summary

Five methods of teaching the understanding of Katona's card tricks were used with high school seniors. Three of the methods required *S* to prepare diagrams (Katona's method and two variations of it described as the horizontal row and vertical column methods). The other methods introduced novel approaches, one using slips of paper as intermediaries, the other consisting of working backwards. Thirty *S*s, 15 males and 15 females, performed by each method.

1. The over-all differences in success according to method were slight. At the point of greatest difficulty, where for the first time *S* was required to skip two cards instead of one, Katona's original method and the slips of paper method proved superior to the other three methods.

2. Analysis of errors on this most difficult transfer task showed the most common error (63 of 109 errors) to be directly related to the requirement to skip two cards instead of one. This is in part a perseverative error, some *S*s continuing to skip one card throughout, as learned in the earlier tasks. Some errors were unique to the individual methods and depended on mechanical features of the method. For example, the slips of paper method permitted the unique error of faulty labeling. The prevalence of careless errors suggests that it is important in teaching understanding to devise methods that are open for review and check through the retracing of steps.

3. Analysis of the improvement between the most difficult task and the one that followed it furthered the impression that casual circumstances of method and attitudes of caution or carelessness were more important in determining the level of error than lack of understanding of the tasks.

4. The superiority in performance of the boys compared to the girls was significant at the .001 level.

References

1. Hilgard, E. R., Irvine, R. P., & Whipple, J. E. Rote memorization, understanding, and transfer: an extension of Katona's card-trick experiments. *J. exp. Psychol.*, 1953, **46**, 288–292.
2. Katona, G. *Organizing and memorizing.* New York: Columbia Univ. Press, 1940.
3. McNemar, O. W. *Word association, methods of deduction and induction, and reactions to set in good and poor reasoners.* Stanford, Calif.: Dept. of Psychol., Stanford Univ., 1954. (Tech. Rep. No. 2, Contract Nonr 225 [02].) (To appear also in *Amer. J. Psychol.*)
4. McNemar, Q. *Psychological statistics.* New York: Wiley, 1949.
5. Piaget, J. *The psychology of intelligence.* New York: Harcourt, Brace, 1950.
6. Sweeney, E. J. *Sex differences in problem solving.* Stanford, Calif.: Dept. of Psychol., Stanford Univ., 1953. (Tech. Rep. No. 1, Contract N6onr 25125.)

(Received for early publication February 11, 1954)

Journal of Experimental Psychology
Vol. 47, No. 6, 1954

LEARNING A RATE OF MOVEMENT[1]

ROBERT S. LINCOLN

The Johns Hopkins University

The production of different rates of movement is a common occurrence in everyday life. Skill in numerous activities often requires both the ability to produce the necessary movements and the ability to produce the movements at particular speeds. It is likely that the acquisition of skill depends, in part, upon learning to make movements at the appropriate rates.

Three problems are considered in this experiment. (*a*) How efficiently can people learn to produce a rate of movement? (*b*) Is the efficiency of learning related to the type of error information provided? Will a continuous source of error information be more useful than an intermittent source? (*c*) Can a rate of movement be maintained on kinesthetic cues alone?

METHOD

Experimental Conditions

The experiment consisted of four successive sessions: (*a*) the prelearning period, (*b*) the learning period, (*c*) the criterion period, and (*d*) the relearning period.

Prelearning period.—Fifty *S*s were randomly assigned to five groups. The *S*s in all groups were required to produce a given rate of movement by turning a handwheel in a clockwise direction. The rate was described to *S*s as a number of units of distance to be traveled in 1 min. All *S*s were to produce the same rate, but the instructions given to the groups differed in regard to the units of measurement with which the rate was described. The *S*s were told to produce a rate either of 100 rpm (angular rate), 160 ft. per min., 1,900 in. per min., or 4,800 cm. per min. (linear rates). Two groups of ten *S*s each worked to produce the rate of 160 ft. per min. All rates were essentially equivalent because the same handwheel was used by all groups. The *S*s were allowed five trials, each 15 sec. long. During the prelearning trials, *S*s did not know how their produced rates differed from the required rate.

Learning period.—In the learning trials, *S*s were told the direction and amount of their average rate error. The amount of error was described in the same units that appeared in their original instructions. Information was supplied in steps of 2 rpm or in appropriate linear extents. The *S*s were to attempt to produce the standard rate by making use of this error information. The learning session included 15 trials, each 15 sec. long. Trials were separated by an interval of 30 sec. The *S*s received their verbal-error information at the end of the intertrial interval. During the learning period one of the groups that had been attempting the rate of 160 ft. per min. was also provided with a continuous visual indication of rate errors in addition to the information about average error.

Criterion period.—Following the learning trials, all *S*s were required to operate for ten additional trials with no information about their rate errors. This was the criterion period. These trials served as a means of evaluating the degree to which the rate of movement had been learned and retained.

Relearning period.—After the criterion period, correction information was again supplied in a series of ten relearning trials.

Apparatus and Procedures

For the purposes of this experiment a device was used that permitted the measurement of the accuracy of a rate of rotary movement. Rotary movements are convenient because they can be maintained for long periods of time without reaching the limit of extent of arm motion.

The apparatus has been previously described in detail (2). The *S*s were seated in a chair. Their task was to turn a handwheel that was located for convenient operation with the right hand. All *S*s in this experiment were right-handed males. The radius of the handwheel was 7.5 cm.

A continuous visual indication of error in rate

[1] This research was done under Contract N5-ori-166, Task Order I, between the Office of Naval Research and The Johns Hopkins University. This is report No. 166-I-181, Project Designation No. NR 145-089, under that contract. Miss Frances Wolfram aided in the collection and analysis of the data.

of turning was provided for certain *Ss*. This indication was a small spot of light that moved to the right of a stationary target when the rate of turning was too high and moved to the left of the target when the rate of turning was too low. The magnitude of the error was reflected by the size of the discrepancy between the positions of the light spot and the target.

Three measures of performance were obtained. (*a*) The total time during each trial that the rate of movement was maintained within tolerance limits of ±9% was recorded on an electric clock. This was a measure of rate accuracy. The *Ss* were not aware that there was any tolerance in the scoring system. (*b*) The actual number of revolutions that the handwheel turned was recorded by an electric counter. (*c*) A second electric counter recorded the number of times that changes in rate exceeded the tolerance limits. This was a measure of the number of oscillations in rate within a trial period.

Trials were automatically timed with a clock that stopped the recording equipment at the end of each trial. Those *Ss* who were required to operate without the visual indication were permitted an adjustment period of 5 sec. before recording began. Recording for the visually aided *Ss* began when they first achieved the required rate of movement. This procedure may have favored the visually aided group to some extent because, for that group, error recording always began at a time when *Ss* were producing the required rate. For the other groups recording began at the end of a fixed interval regardless of the rate that was being produced at the time. When visually aided *Ss* were not receiving visual information, they were also allowed a 5-sec. adjustment period.

The *Ss* knew that the trial period would be less than 1 min. long. They were told that the required rate of cranking was well within their physical capabilities. The *Ss* were cautioned not to attempt to count the number of revolutions while turning the handwheel. They were required to keep their eyes closed during each trial, except when the visual indication was provided.

Results

Several analyses of variance were performed on the data. Error terms for both correlated and noncorrelated scores were used where appropriate. For all analyses the homogeneity of the variance estimates was tested by means of Bartlett's test. In one analysis the variance estimates were found to be heterogeneous. The data involved were, therefore, reanalyzed by the nonparametric analysis of variance described by Friedman (1). Differences were accepted as significant if the p values were equal to, or less than, .05.

Prelearning period.—The initial level of rate accuracy was established during the prelearning trials. Different rate descriptions were used because it was thought that accuracy might be related to the magnitude of the numbers involved in the descriptions.

There was a tendency for the constant error in the production of the rate to increase with the magnitude of the numbers involved in the linear rate descriptions, but the tendency was not significant. In fact, the mean square for error was larger than the mean square for rate descriptions. Apparently the task was an extremely ambiguous one. Those *Ss* who worked with the angular description showed the greatest constant error, although the number in their instructions was the smallest.

In general the direction of the constant errors was positive, and there was a significant tendency for the rate of turning to increase as trials progressed.

Learning period.—When error information was provided, all groups showed rapid improvement in rate accuracy. Maximum accuracy was reached in a very few trials.

This result is pictured on the left side of Fig. 1. In Fig. 1, Trial 5 is the fifth prelearning trial. Accuracy was measured as the number of seconds during a trial in which the rate of movement fell within ±9% of the required rate.

The type of rate description was a significant source of variation during the learning period. Accuracy appears to be inversely related to the

magnitude of the numbers provided in the error information. This result, together with the evidence that these *Ss* were poor interpreters of verbally described rates, strongly suggests that *Ss* were merely adjusting their output according to the magnitudes of the numbers with which they worked. It is likely that the same results could have been obtained if numbers alone had been used to describe the errors. The fact that the numbers were coupled with units of measurement probably had little to do with *Ss'* performances.

As measured in this experiment, accuracy scores are made up of two components. When the actual rate of turning differs greatly from the required rate, accuracy is largely determined by the size of the difference. When the difference between the actual and the required rate is small, accuracy is largely determined by the steadiness with which the rate is maintained within tolerance limits.

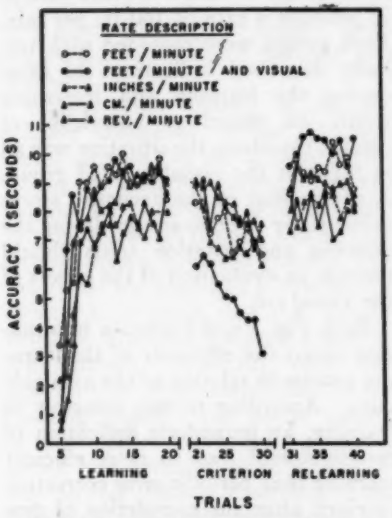


FIG. 1. Rate accuracy obtained with various rate descriptions under three different experimental conditions. The ordinate shows the number of seconds for each trial that the rate of movement was kept within ±9% of the standard rate. The maximum possible score is 15 sec.

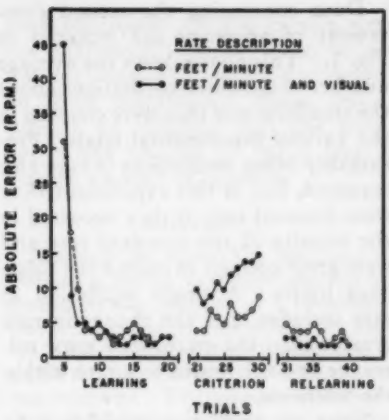


FIG. 2. Average error in rate obtained with and without the visual indication of error. Error was measured as the mean of all discrepancies between standard and actual rate summed without regard for sign.

It is possible to obtain an idea of what *Ss* were learning by inspecting curves that separate out the two components of accuracy. Figure 2 pictures the absolute rate error produced in the various experimental periods. Error was measured as the mean of all discrepancies between the standard and the actual rate summed without regard for sign. Again Trial 5 is the last of the prelearning trials. Results for only two groups are shown to improve readability. Data for the other groups are highly similar. From Fig. 2 it is obvious that *Ss* did not initially know how to produce the required rate. As soon as error information was provided, however, the verbally aided *Ss* were quickly able to align their rate production with the number system. The visually aided *Ss* were also able to adjust quickly to the required rate.

Data concerning the second component of accuracy are pictured in Fig. 3. This figure shows the average number of measured oscillations about the standard rate that were counted in the various experimental trials. Presumably other oscillations in rate also occurred, but in this experiment they were counted only if they occurred in the vicinity of the standard rate and were great enough to exceed the tolerance limits. A single oscillation in rate includes both the change in rate greater than the established error tolerance and the return to a rate within the tolerance.

Since an oscillation could not be counted unless it occurred near the required rate, the early part of the learning curve in Fig. 3 merely reflects the fact that *Ss* were approaching the standard rate as trials progressed. By Trial 12, however, the size of the absolute error in Fig. 2 had reached a relatively steady level. With absolute error constant, it becomes reasonable to look for trends in the frequency of rate oscillations. Improvement in performance would be shown by a reduction in the number of oscillations in rate. Inspection of Fig. 3 indicates that there was no consistent tendency for the oscillations to decrease with practice either in the original learning period or in the relearning period. It appears that *Ss* did not learn to eliminate their rate oscillations. This is true even of the group that received visual information regarding the variability of their performance within a trial.

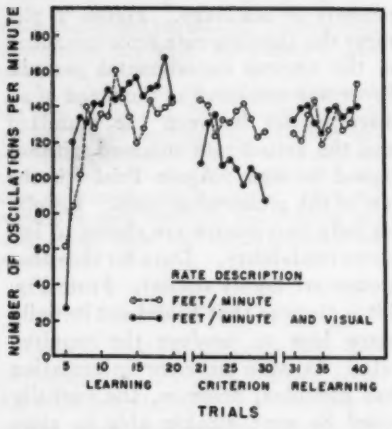


FIG. 3. Rate oscillations made with and without the visual indication of error. A single oscillation includes both the change in rate greater than the established error tolerance and the return to a rate within the tolerance limits.

Previous work (2) has suggested that *Ss* are unable to use all of the information provided by a continuous visual indication of rate error in a handwheel cranking task. The visually aided group was included in the present experiment to help determine whether or not this aid was of value. Both the visually aided group and one other group worked with instructions to produce a rate of 160 ft. per min. Both groups were provided with verbally described indications of error during the learning and relearning trials. In regard to knowledge of results, therefore, the situation was all in favor of the visually aided group. A comparison of the accuracy scores achieved by the two groups during the learning and criterion trials should provide an evaluation of the effects of the visual aid.

Both Fig. 1 and 2 contain information about the efficiency of the learning process in relation to the available cues. According to one principle of learning, an immediate indication of results should lead to more efficient learning than periodic error correction received after the completion of performance. In Fig. 1 and 2, such an effect is apparent only in the first few trials. Those *Ss* who were supplied with the visual error indication showed the highest rate of improvement early

in the learning period. This advantage did not continue. Later in the learning period the visual aid was of no value. The accuracy of the comparable verbally instructed group was equal to the accuracy achieved by the group that received the visual indication as well as the verbal corrections.

Criterion period.—When error information was no longer provided, *Ss* were forced to rely exclusively on kinesthetic cues in the production of the required rate of movement. The attempt was made, during the criterion period, to determine the degree to which the rate of movement had been learned and retained.

The center curves in Fig. 1 and 2 show that all groups had learned the rate of movement to some extent. Accuracy was much greater in the criterion trials than in the last prelearning trial (Trial 5). The systematic decrease in the accuracy obtained was significant, but a longer training period might have reduced this tendency.

Of particular interest is the significant difference in the accuracy scores obtained by the various groups during the criterion period. The visual aid had been only of slight value in the learning trials. During the criterion trials, previous practice with the visual aid produced a detrimental effect. It is likely that the visually aided *Ss* depended to a great extent on the obvious, dominant cues and ignored the verbal corrections in the learning period. The poorer quality of the verbal corrections probably forced the verbally aided groups to rely on less obvious kinesthetic cues. Because of this, their performance was relatively better when all *Ss* were forced to rely exclusively on kinesthetic cues during the criterion trials.

Relearning period.—In the relearning period the differences between the accuracy scores for the groups were not significant. Furthermore, no significant change in accuracy appeared as trials progressed. These results provide no evidence that the visual cue might become of value after a greater amount of practice.

Discussion

Learning the rate.—The efficiency of learning a rate of movement was not improved by the addition of a continuous visual indication of rate error. In fact, the visual aid reduced the extent to which the rate was retained when the cue was removed. The amount of transfer from the learning task to the criterion task was less for the group that received the visual aid than for the groups that learned without it. These results have general implications for the design of training tasks from which positive transfer effects are expected. The inclusion of extra cues not found in the criterion task may be a hindrance, rather than an aid, when performance on a criterion task is required.

Rate errors in tracking performance.—After the final level of rate accuracy was reached in this experiment, the frequency of measured rate oscillations remained relatively constant at about 140 per min. This was true for both the visually aided and the comparable verbally aided groups. The visual aid was of no value in reducing this oscillatory feature of the movement. This fact suggests that most of these oscillatory errors are not intentionally corrected by the human operator, but are inherent in the nature of rotary motion.

Previous work (2) has shown that the frequency of oscillatory errors increases directly with the rate of rotary movement. This result also supports the idea that the oscillatory errors are automatically generated and corrected in the process of producing a rate of rotary movement. Such movement involves rapid changes in the direction of force, with and without the aid of gravity. The most probable assumption is that the oscillatory errors are often generated in one part of the

rotational cycle and corrected in a succeeding part of the cycle as a change in the direction of force occurs.

Attempts to uncover periodicities in tracking records generated with a handwheel control must take into account the high frequency errors produced in rotary movement. Vince (1) has reported the occurrence of corrective movements at a rate of 2 per sec. in a handwheel tracking task. In this task, however, a handwheel speed of 1 rpm was used. At such a slow rate the properties of a handwheel are quite different from those that appear at higher rates of movement.

Summary

This experiment was concerned with the ability of human *Ss* to learn a rate of movement. The rate of movement was produced by turning a handwheel. The rate to be produced was described to different groups of *Ss* as various units of distance to be traveled in 1 min. of time. After initial attempts to produce the rate, all *Ss* received verbal descriptions of their average rate error. This information was to be used in the next attempt at producing the rate. At this time one group was also provided with a continuous visual indication of rate error as an aid in learning the rate. Retention of the rate of movement was evaluated in a series of criterion trials by requiring *Ss* to produce the rate without further error information. The *Ss* then relearned the rate of movement.

Before error information was provided, *Ss* tended to crank at a rate higher than the required rate. This tendency increased as trials progressed. The effect of the different rate descriptions was not significant.

When correction information was provided, all groups showed a rapid increase in rate accuracy. The *Ss* who received the same type of verbal corrections achieved as high a level of accuracy as those *Ss* who were also provided with the visual aid.

The *Ss* were able to learn and retain the rate of movement. In the criterion trials the rate was produced with much greater accuracy than that obtained in the prelearning period. The group that worked with the visual aid showed the poorest rate accuracy when forced to depend upon kinesthetic cues. The extra knowledge of results provided by the visual aid was a detriment during the criterion trials.

These results appear to have implications for the design of training equipment. In addition, the results also provide evidence that rate accuracy in handwheel cranking is limited by the nature of rotary movements.

References

1. Friedman, M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. statist. Ass.*, 1937, **32**, 675–701.
2. Lincoln, R. S. Rate accuracy in handwheel cranking. *J. appl. Psychol.*, in press.
3. Vince, M. A. The intermittency of control movements and the psychological refractory period. *Brit. J. Psychol.*, 1948, **38**, 149–157.

(Received for early publication March 4, 1954)

LEARNING THEORY, PERSONALITY THEORY, and CLINICAL RESEARCH

The Kentucky Symposium . . .

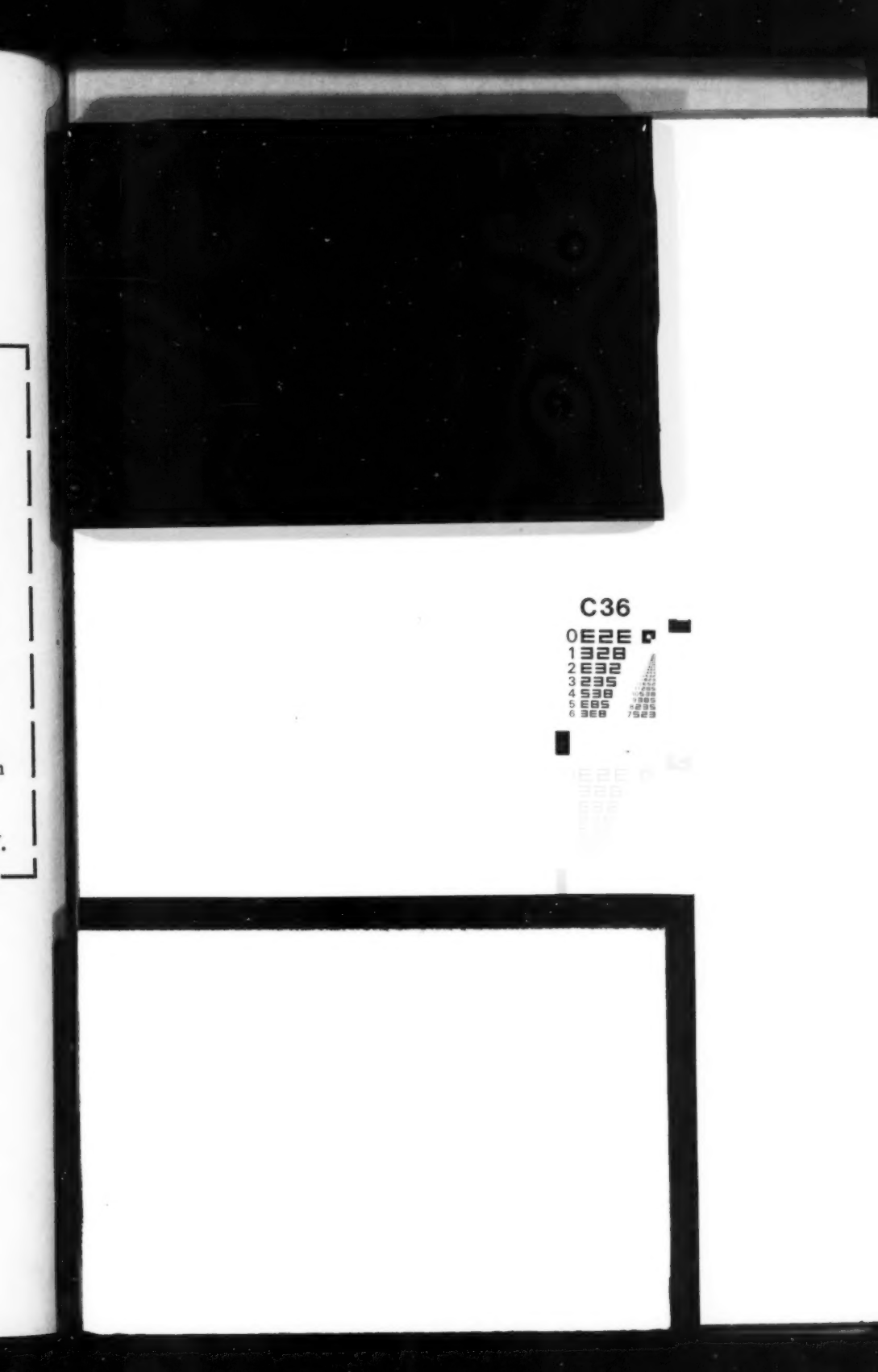
This book contains eleven papers encouraging a closer integration of three branches of psychology—learning theory, personality theory, and clinical research. Each report is the work of a specialist writing in one or more of the three areas. Presented at a psychological symposium held at the University of Kentucky in 1953—the collected lectures provide a comparison and an extremely interesting account of the most recent thinking in these fields. Not mere summaries of scattered thought, the papers point the way future researches will take into many as yet unanswered problems. Here is a brilliant multiple approach to learning theory, personality theory, and clinical research, straightforwardly conceived, written, and edited. *1954. 164 pages. $3.50.*

Contributors . . .

D. K. Adams
R. B. Ammons
J. M. Butler
R. B. Cattell
H. F. Harlow
N. R. F. Maier
O. H. Mowrer
D. Snygg
K. W. Spence
D. D. Wickens
J. R. Wittenborn

Send for an on-approval copy

JOHN WILEY & SONS, Inc., 440-4th Ave., New York 16, N.Y.


C36
0E2E
1328
2E32
3235
4538
5685
6388